Rhythm & Cues

Rhythmic structure and segmentation
in early language acquisition

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: *"Letters and numbers attracted by a spark III"* by Joan Miró.

ISBN: 978-94-6093-233-5
NUR 616

Rhythm & Cues
Rhythmic structure and segmentation
in early language acquisition

Rhythm & Cues
Ritmische structuur en segmentatie
in vroege taalverwerving

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
vrijdag 31 maart 2017
des middags te 2.30 uur

door

Brigitta Marinka Keij

geboren 30 juni 1986
te Voorburg, Nederland

Promotoren: Prof. dr. R.W.J. Kager
Prof. dr. W. Zonneveld

The research reported here was part of the project 'Parsing and Metrical Structure: Where Phonology Meets Processing' supported by the Netherlands Organisation for Scientific Research under project number 360-89-030.

Contents

List of Tables

List of Figures

Acknowledgements

"Feeling gratitude and not expressing it is like wrapping a present and not giving it." –*William Arthur Ward*

I hope I have not wrapped a present for someone without giving it, but in case you feel forgotten, please imagine your name in writing below.

First of all, I would like to thank my supervisors René Kager and Wim Zonneveld. René, thank you for believing in me from the start of the project until the very end of writing this dissertation, even though it took a bit longer than probably both of us anticipated on. You were always optimistic and available to help, which is amazing considering the amount of tasks on your plate. Your knowledge on many different topics sometimes intimidated me, but I also knew you would never judge my lack of knowledge in some of these areas. Wim, thank you for your good sense of timing and showing up exactly at the right moments to steer my thinking and writing in the right direction. Even after you retired, you were still available as the steersman of the project, which I appreciate very much.

Secondly, I would like to thank my project-buddy and paranymph, Sandrien van Ommen. Thank you for being different from me in many ways, but also sharing many of the same values both in science and in life in general. I think we learned a lot from each other over the past years, as researchers, but also definitely as friends. You were always there for me, in good times and in less good times, scientifically as well as privately. (Gaaike, thank you for the title of my dissertation!) I would also like to thank my other project-mates, Anna Kijak and Tom Lentz. Anna, thank you for being there for the kick-off of the project and providing me with many cues to launch into the topic of metrical segmentation. Tom, thank you for taking over Anna's role and for your welcome expertise on eye tracking and statistics. One of your expressions I will never forget: absence of proof does not entail proof of absence.

Thirdly, I would like to thank Paula Fikkert and Bariş Kabak for being both part of the advice committee of the project and of my reading committee. I would also like to thank Laura Bosch, Thierry Nazzi and Frank Wijnen for being on my reading committee. (Frank, thank you for supporting me in many ways!)

I would like to express a special thank you to all the parents and babies that participated in my experiments, both in Utrecht, the Netherlands, and in Istanbul, Turkey. It is quite something to participate in experimental research with a tiny baby of only a few months old, but still these parents were willing to spend their time and energy on coming to the lab, for which I am very grateful. Without them this dissertation would not have been here today.

Next, I would like to thank all the colleagues who helped me collect the data for this study in Utrecht as well as in Istanbul. First of all, I would like to thank Aylin Küntay for agreeing to collaborate in this project and facilitating the use of the equipment and infrastructure of the Language and Communication Development Lab in the Psychology department of Koç University in Istanbul. Secondly, I would like to thank Şule Kurtçebe for being my savior: without you I would have never been able to test all those Turkish-learning babies. Thank you very much for your open mind, flexible attitude, work ethic and warm personality. I enjoyed working with you very much. I would also like to thank Özge Savaş, Gizem Köksal and Ebru Ger for being so collaborative and persistent in recruiting and testing Turkish-learning babies in Istanbul. For technical support in Istanbul I would like to thank Mustafa Kaya. In Utrecht, I received lots of technical help from Alexis Dimitriadis, Martijn van der Klis, Alex Manus, Iris Mulders, Chris van Run and Theo Veenker. Lianne Dekker helped me a lot in the final stages of my project: thank you for being eager to learn and being a very reliable testing assistant. For help with the statistical analyses, I would like to thank Huub and Mattis van den Bergh, like father, like son!

There are also many colleagues who were not involved in my project directly, but nonetheless supported me during my PhD research. Firstly, I would like to thank my long term officemates Lizet van Ewijk, Eva Poortman, Yipu Wei and Rob Zwitserlood. You made me feel right at home in the Trans building and we shared the same office routines, which I enjoyed very much. Thank you for having such a 'gezellige tijd' with me! I also enjoyed the fact that Lizet and Rob worked with adults and children with language disorders, so I could talk about these topics and stayed up to date, even though they were not part of my own PhD research. Thank you for keeping me involved in this field of research! Eva, thank you for being the visual stimulus in most of my experiments! Yipu, thank you for your kindness and attentiveness! Secondly, I would like to thank my shorter term officemates, Marta Castella and Marko Hladnik, and my post-PhD-project officemates Frans Adriaans, Hugo Schnack and Hayo Terband for their company during office hours.

Outside the office there were also many work-related and not-so-work-related meetings and gatherings with my Utrecht colleagues which either contributed to or hindered my work (by having too much fun), respectively. The work-related meetings were with colleagues working on similar topics, such as the Experimental Phonology Group meeting and the Babylab meetings. Thank you for all your input: Antonio Baroni, Hans Rutger Bosker, Jeroen

Breteler, Jennifer Cabrelli Amaro, Desiree Capel, Ao Chen, Aoju Chen, Ileana Grama, Nadya Goldberg, Mengru Han, Sander van der Harst, Carolien van den Hazelkamp, Shuangshuang Hu, Nivja de Jong, Annemarie Kerkhoff, Maartje de Klerk, Arnout Koornneef, Liquan Liu, Zenghui Liu, Giorgio Magri, Anne-France Pinget, Silvia Radulescu, Anna Sara Romøren, Hans Van de Velde, Anqi Yang and Jingwei Zhang. The not-so-work-related gatherings often included the people mentioned above, but also took place in the good company of: Stavroula Alexandropoulou, Daria Bahtina-Jantsikene, Dominique Blok, Xiaoli Dong, Marjolein van Egmond, Anja Goldschmidt, Alexia Guerra Rivera, Mirjam Hachem, Bjorn 't Hart, Lotte Hendriks, Heidi Klockmann (thanks for proofreading!), Misha Knyazev, Loes Koring, Bert Le Bruyn, Anne van Leeuwen, Sophia Manika, Hannah De Mulder, Liv Persson, Jolien Scholten, Maartje Schulpen, Marko Simonović, Marijn Struiksma, Assaf Toledo, Anna Volkova, Hanna de Vries and Franca Wesseling. Thank you for all the fun we had!

Outside the Utrecht Institute of Linguistics OTS, I also met many new colleagues who contributed to my PhD work, for instance at the Utrecht-Nijmegen meetings. Thank you for your input as well: Titia Benders, Christina Bergmann, Helen Buckler, Nienke Dijkstra, Imme Lammertink, Stefanie Ramachers, Tineke Snijders, Antje Stöhr, Sho Tsuji and Maarten Versteegh. We also had national babylab meetings, called the Baby Circle, at which I met colleagues from Amsterdam, Leiden, Nijmegen and Utrecht who I would like to thank for their input: Elma Blom, Carlijn van den Boomen, Andreea Geambaşu, Sabine Hunnius, Caroline Junge, Claartje Levelt and Josje Verhagen. I also met a lot of new colleagues from outside the Netherlands during the LOT schools and all the conferences and workshops abroad I was lucky enough to be able attend and learn from. I learned both from the people I met and from traveling to many amazing places, thank you all!

Besides all the 'frolleagues' mentioned above, there are of course also many non-work-related friends that have supported me throughout this endeavor. I will thank them in chronological order of when we met, because I would not know where to start otherwise. Renée, you are officially my oldest friend (for 18 years!) and one that I can always rely on, for better or for worse and with many laughs and sometimes tears (also often from laughing, by the way). Thank you for sticking with me for so long! Igor, we have also known each other for a very long time and I have always valued our friendship deeply, even though we do not see each other as often as we would like. Thank you for being my philosophical mirror. Irene and Loes, we met during our BA programme before we knew we would each go in a different direction, but nonetheless we have managed to keep our friendship alive, for which I am very thankful. Suzan, my best Utrecht friend and other paranyphm. Linguistics is what brought us together, but we soon discovered that we shared so many other interests: surfing, longboarding, snowboarding and watching *Friends*. You made and still make my life in Utrecht so much more fun! Thank you as well for embarking on the PhD adventure together and almost defending at the same time! Julia, Kika and Vanessa, I am so happy to have met you during our international master programme! You are my furthest friends, but at the same time among those closest to my heart and I am sure you will stay there forever. Kika and Vanessa, also many thanks for taking the effort to proofread my dissertation! Eliane and Roos, you were there almost from the start of my PhD project when we lived in a shared house. You were my distraction from the PhD stress, which you relieved with many 'tassen' tea and watching *Embarrassing Bodies*. Thank you for being there and for being my home.

A special thank you goes out to the people who helped me along the road towards becoming a PhD candidate from very early on. There have been many intersections along this way at which I might have taken a different direction if it was not for them: my teachers in primary school for not underestimating me despite my slow work pace, my teachers in secondary school for advising me to go to VWO-level even though I preferred to go to HAVO-level with my best friends, my aunt, Tante Carrie, for questioning me when I considered starting a BA at a university of applied sciences, my first serious employer and academic mentor, Karijn Helsloot, for nourishing my decision to continue my studies with an MA programme, and the supervisors of my first master thesis, Leonie Cornips, Roeland van Hout and Aafke Hulk, who encouraged me to submit my work to an international conference and a scientific journal. All of you influenced the outcome of the decisions that ultimately led to applying for the PhD position which resulted in this dissertation. In this context I would also like to thank my PhD mentor, Simone Veld, for all your advice on planning, time management and even winter coats (!) when I needed it the most. Without your help this dissertation would have been here as well, but the road towards it would have been much harder.

All of this studying would not have been possible without the literal and figurative support of my family. Pap en mam, dankzij jullie heb ik altijd mijn hart kunnen volgen in mijn studiekeuzes en was er meer dan voldoende ruimte om mezelf academisch te ontwikkelen, duizendmaal dank daarvoor! Pap, hoewel je niet altijd geïnteresseerd was in de inhoud van mijn studie en PhD project, was je wel altijd geïnteresseerd in het verloop ervan en je hoge verwachtingen hebben me uiteindelijk gebracht waar ik nu ben. Mam, ik vind het heel leuk dat wij onze interesse voor het taal-lerend kind delen en dat je zelfs mijn eerste masterscriptie helemaal hebt proefgelezen en nu mijn Nederlandse samenvatting. Bedankt voor alle steun! Kir, mijn grote zus, een rol

die je met verve vervult. Jij ging me altijd voor in het doen van spannende nieuwe dingen, zoals voor langere tijd naar het buitenland voor studie of werk, en daardoor durfde ik ze ook aan. Met trouwen en kinderen krijgen ben je me ook voor, maar wat ben ik een trotse tante! Sjoerd, Mika en Jari, bedankt voor de leuke momenten die we samen beleefd hebben. Ik kijk uit naar alle momenten die nog volgen in de toekomst!

Naast mijn eigen familie wil ik ook graag de familie van Matthijs bedanken, omdat zij eigenlijk ook als mijn familie voelen. Eduard, bedankt voor je warmte en ruimdenkendheid, bij jou kan ik altijd helemaal mezelf zijn en dat vind ik best bijzonder. Annemarieke en Aad, bedankt voor jullie gezelligheid tijdens etentjes, dagjes en weekendjes weg! Elise, Jan, Sabine en Marijn, ook jullie wil ik heel erg bedanken voor de Brabantse gezelligheid en alle leuke momenten samen! Jelmer en Melina, for you I am switching back to English: thank you for all our shared dinners and evenings in Utrecht, which will hopefully be followed by many more, maybe in Amsterdam and Geneva! Voor alle oma's aan wie ik dit boek beloofd had, maar voor wie het helaas te laat kwam, bedankt voor jullie (eigen)wijsheid waar ik heel veel van geleerd heb.

Tot slot, de persoon met wie ik inmiddels mijn halve leven al heb gedeeld, Matthijs. Jij houdt me met beide benen in de echte wereld als ik weleens wegdroom naar mijn eigen wereld, maar jouw grote dromen inspireren mij ook om ambitieus te zijn. Met jou geniet ik van de kleine dingen in het leven door ze groots aan te pakken. Ik houd van jou, lief!

Chapter 1
General introduction

Adult speakers have unconscious knowledge about the sound regularities of their language(s). These sound regularities are part of the phonology of a language and include knowledge on the rhythmic structure of a language. Rhythmic structure, as described in metrical phonology, determines, among other aspects, where stressed syllables occur in words. In many languages, stress falls near the edges of words, most typically on the initial, final or prefinal syllable. It has been hypothesized that this delimitative property of word stress may be one of the phonological properties of words that assists listeners in parsing the continuous speech stream into words. For example, in an initial stress language, assuming one-and-only-one main stress per word, every stressed syllable can reliably be taken to mark the beginning of a new word. However, while both left and right edge stress languages exist, experimental evidence for this strategy labeled the Metrical Segmentation Hypothesis (Cutler & Norris, 1988) comes almost exclusively from initial stress languages, while metrical segmentation in non-initial stress languages was hardly investigated until recently (see van Ommen, 2016). This dissertation will add an acquisition perspective on language-specific parsing by taking a cross-linguistic approach to investigate the role of rhythmic structure for speech segmentation in early language acquisition. By adopting a cross-linguistic perspective, this study provides an answer to the question whether learners of non-initial stress languages also use cues that specifically fit the metrical structure of their native language.

Perhaps the most fundamental question of phonology is why there is phonology at all. Why do speakers not simply utter individual sounds in a row? Why do they combine them into syllables and put prosodic patterns on them? At least a partial answer to these questions has in the past few decades been formulated by a cognitive view on phonology (Kaye, 1989). By recognizing that phonology is not only about uttering speech sounds ('ease of articulation') but also about mental representations of sound sequences, it is now widely accepted that the brain contains a lexicon in which the phonological forms of word-like units are stored and that these representations are used for processing speech efficiently. These units, for current purposes 'words', may be affected by phonological processes that typically occur at this word level and are distinct from phrasal phonology, which typically affects word sequences (Chomsky & Halle, 1968, Selkirk, 1980; Kiparsky, 1982). The aim of this dissertation is to contribute to a better understanding of how the phonological representations of these word-like units are acquired, and in particular, how the word level phonological phenomenon of stress is acquired as part of these word representations.

Harking back to the early work of Trubetzkoy (1939/1969), word stress manifested itself as a boundary phenomenon in the metrical phonology of Hayes (1980) and many others by so-called End Rules, as well as in Optimality Theory by McCarthy & Prince (1993), Kager (1997), Pater (2000) and others by using a theory of alignment. Theoretical notions such as these will only be implicit in the current dissertation. More pertinently, the focus of this dissertation is on acquisition aspects of the research area described: it addresses the nature of young listeners' mental representations of word stress and how these representations may contribute towards parsing, or segmenting, a string of sounds into words.

There are two issues concerned with investigating how these mental representations, potentially, function as boundary signals for infant learners. The first involves the balance between language-specific cues and universal cues for speech segmentation. To what extent are young listeners, say those of an age up to 8 months, already guided by knowledge specific to their native language? Or, conversely, how much of metrical parsing is universal, somehow 'inborn' or independent of experience with the native language? Answering these questions is of importance for our understanding of language processing in general, that is, how much processing is guided by universal biases not acquired from the input? The second issue concerns the representation of word stress involved in parsing: do young listeners exclusively represent the acoustic properties, such as pitch, duration and intensity, that signal word boundaries, or do they parse speech using abstract metrical units by mentally imposing rhythmic grouping on syllables?

Providing answers to these questions will clarify the relation between the learning system and the mental representations that are part of a learner's linguistic knowledge. Word stress is especially suitable for this type of study for two reasons. First, because of its abstract nature: adult humans perceive a rhythmic organization in syllable sequences even in the absence of the relevant acoustic information (Schreuder, 2006). The expectation, for that reason, is that metrical parsing involves abstract metrical representations, not just low-level acoustic information. Important to note here is that these two issues are mutually independent. If it were, for example, found that young listeners parse speech using abstract metrical units, these units might either be part of their language-specific knowledge or be universally available. Whatever the nature of these mental representations, the central question will be how they are internalized and, with that, to which degree these representations depend on linguistic experience. Investigating language acquisition allows assessing this balance between language-specific and universal parsing biases, as well as assessing how abstract metrical representations develop.

Second, word stress is an excellent candidate for cross-linguistic research because stress systems differ typologically along dimensions of which a substantial amount of understanding has been gained. A cross-linguistic perspective also makes it possible to compare universal and language-specific rhythmic preferences. The default position (i.e. preferred position from the perspective of the language user) of stress within words differs between languages and covers the three most common word positions in the stress typology mentioned above: initial, final and prefinal. A cross-linguistic study on word stress would ideally examine languages with typologically different stress patterns. Two languages which meet the requirement of filling different typological slots and which were comparatively easily accessible in the research context of this study are Dutch, which has prefinal, or statistically speaking initial stress, and Turkish, which has default final stress.

To assess how rhythmic sensitivity develops initially, it is essential to go back to the earliest auditory, or even sensory, experiences infants have. Even before the auditory system is functional at the beginning of the third trimester of pregnancy, the fetus has had the sensory experience of the rhythmicity of the mother's heartbeat, its own heartbeat and, also, the rhythmic pattern of the mother walking. Around the 26th week of gestation, the first auditory stimuli are perceived, of which the mother's voice, as well as the speech of other speakers around the mother, is an ever-present source. As the womb functions as a low-pass filter, mostly the rhythmic patterns of speech are conducted and perceptible in utero (see Giovanelli, Sansavini & Farneti, 1999 for a review). This rhythmic patterning is thus the very first source of linguistic input to the unborn infant. In language, rhythm is hierarchically structured and, among other ways, can be established by linguistic stress. Alternating stressed and unstressed elements shape the rhythm of speech at different levels of the hierarchy: at the phrase level and at the word level.

Phrases in spoken language, even when not separated by short pauses, are cued by universal means such as final lengthening, pitch movement and final devoicing, thereby indicating phrase boundaries more consistently across languages than phrase-medial word boundaries. However, as phrase boundaries necessarily align with word boundaries, phonetic properties of the former can provide the infant with the first clues to the rhythmic pattern of words in their native language (Aslin, Woodward, Lamendola & Bever, 1996). As rhythm at the word level is often shaped by the position of the main stress in relation to the word boundary, this word well-formedness can be represented by either a disyllabic strong-weak (SW) pattern in penultimate and initial stress languages, or weak-strong (WS) pattern in final stress languages, depending on the language the infant is hearing. In this dissertation, when I use the term rhythmic pattern (in contrast to 'rhythm class') I refer to word stress patterns that can be represented metrically by a grid such as that depicted in Figure 1.1. The general goal of this dissertation, also reflected in the issues described above, is to determine when infants discover the dominant stress pattern of their native language and whether they subsequently also use this stress pattern for segmenting the speech stream into words.

```
*                   *
*    *         *    *
```

Figure 1.1: *Metrical grids representing an SW and a WS word-unit, respectively.*

Regarding the key issues of this dissertation, the literature more or less converges on the view that infants, like adults, may use rhythmic information when learning to parse the continuous speech stream into words, or word-like units, of their language, i.e. for acquiring an ability called word segmentation. Word segmentation is a prerequisite for word learning, as the mapping of form onto meaning can only be done once a phonological form of the word is established. From the rhythmic structure of the language the infants are hearing, they may learn about the dominant (i.e. most frequent pattern from the perspective of the learner) rhythmic pattern of words in the language they are learning. From there, word segmentation may be informed by this hypothesized rhythmically well-formed unit. The aim of this dissertation is to test the language-specificity hypothesis of metrical segmentation for native language acquisition in infants aged 4, 6 and 8 months. The dissertation will investigate the emergence of rhythmic preferences and word segmentation in infants learning metrically opposed languages, namely Dutch, which can be characterized as having initial stress (SW), versus Turkish, which has final stress (WS).

To investigate this issue, I take an experimental approach and thus set up a series of experiments trying to answer the following research questions: (i) when do infants acquire the rhythmic pattern of their native language, considering rhythmically opposing languages (SW versus WS)? Consequently, (ii) do they use the stress pattern of their native language to inform word segmentation or do they use universal cues for word segmentation? Before investigating these questions in further detail, I will first review the existing literature on rhythmic structure, rhythmic discrimination and preference studies and speech segmentation studies in Chapter 2. The goal of this discussion is twofold: to review existing hypotheses relevant to the research questions formulated above, and to show that methodological considerations are of the utmost importance regarding the development of rhythmic preferences and the use of stress cues for word segmentation in infants learning a non-initial stress language. After this, I will discuss the phonological properties of the languages under investigation in this dissertation, namely Dutch and Turkish, in Chapter 3. To address research question (i) stated above, two preference experiments were conducted which I will report on in Chapters 4 and 5. The results of these two experiments will be discussed together in an interim assessment in Chapter 6. The second research question (ii), will be addressed by a word segmentation experiment reported on in Chapter 7. Finally, I will test the results of all three experiments against a small body of hypotheses, those addressed in the literature review and new hypotheses following from my experiments. The final chapter, Chapter 8, provides overall conclusions and suggestions for further research.

Chapter 2
Literature review

2.1 Rhythmic structure

Many natural and human phenomena possess rhythm: the heartbeat of animals, the waves of the ocean, music and spoken language. One could ask why rhythm is present in all of these phenomena. The most plausible answer is that it orders or structures them, as is certainly the case for language. In language, rhythmic structure is provided by prominence relations: weak versus strong, high versus low, long versus short. Over the past decades, phonological research on stress systems has converged on a restrictive representational model, namely metrical phonology, which accounts for typological variation among stress languages by making use of either parameters (Halle & Vergnaud, 1987; Dresher & Kaye, 1990; Hayes, 1995) or constraints (McCarthy & Prince, 1993; Kager, 2007). This model of metrical phonology as well as hypotheses from different theoretical backgrounds, such as the Iambic-Trochaic Law (ITL: Woodrow, 1909; 1911; 1951), the Universal Trochaic Bias (UTB: Allen & Hawkins, 1978; 1980) hypothesis and the Rhythm Class Hypothesis (RCH: Pike, 1945; Abercrombie, 1967; Ladefoged, 1975), will be described in some detail below in Section 2.1, to the extent necessary as a background to this thesis.

Metrical representations have found empirical support from language acquisition studies (a.o. Hochberg, 1988; Fikkert, 1994; Nouveau, 1994). Typically, these early studies provide evidence from a language production perspective. In Section 2.2 below, the literature on the role of rhythmic structure (mostly revolving around metrical SW- and WS-units, but sometimes addressing 'rhythm class') in native language acquisition will be reviewed from a perception, or processing, perspective covering both rhythmic discrimination (a.o. Skoruppa, Pons, Christophe, Bosch, Dupoux, Sebastián-Gallés, Alves Limissuri & Peperkamp (2009); Skoruppa, Pons, Bosch, Christophe, Cabrol & Peperkamp, 2013) and rhythmic preference studies (a.o. Jusczyk, Cutler & Redanz, 1993a); Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993b). The focus then shifts to addressing some methodological issues involved in these type of studies in Section 2.3, before proceeding to the literature on word segmentation in native language acquisition (a.o. Saffran, Aslin & Newport, 1996) and the Metrical Segmentation Hypothesis (MSH: Cutler & Norris, 1988), which will both be covered in the last section of this chapter, Section 2.4.

2.1.1 Metrical phonology

One of the most important insights of metrical phonology is that stress is not a feature on par with segmental features such as [voice] and [back], but rather a metrical representation. The most extensive theoretical approach to stress is that of Hayes (1995). However, the pioneers of his ideas were Liberman (1975), Liberman & Prince (1977) and Selkirk (1980), who stated that stress is the linguistic manifestation of rhythmic structure. This rhythmic structure is mainly manifested in the perception of spoken language, although not necessarily phonetically present in the production of spoken language. This is due to the fact that the human cognitive system tends to structure incoming stimuli in general, even when this structure is not physically present in the stimuli (Schreuder, 2006). Nonetheless, stress is phonetically marked by a combination of at least three, sometimes four, factors: pitch, duration and intensity, and segmental processes, such as vowel reduction.

Metrical phonology originally employed two distinct hierarchical structures: the S/W relational tree (Selkirk, 1980) and the metrical grid (Prince, 1983). Later, these two structures merged into a constituent structure, a so-called 'constituentized grid' (Hammond, 1984; Halle & Vergnaud, 1987; Hayes, 1995), providing the basis for a temporal-rhythmic interpretation based on metrical feet which is of interest to the purpose of this dissertation. The notion of rhythmicity is found to play a central role in determining patterns of prominence in both phrases and words. Prince (1983) argued that phrases and words should be related directly to the grid, without the intervention of a level where calculations with S/W take place on trees. When infelicities in the grid, such as stress clash (two adjacent stressed syllables), appear in the normal course of linguistic concatenation, it is often the case that various steps are taken to repair them, e.g. the Rhythm Rule of English. Prince (1983) discusses clash avoidance (*Clash) and the Perfect Grid (PG: avoiding lapse, two adjacent unstressed syllables) as central causes in the determination of the character of a stress pattern. Clash and PG are the major determinants of alternating patterns: patterns which build a maximal organization into the grid. Neutral between the two theories, the S/W relational tree and the metrical grid, is the assumption that a word must somewhere bear a (primary) stress. Stress is culminative, which means there is at least one prominence peak per word. This peak is determined in the lexical representation of a word on the metrical level, marking phonemic contrasts and morphological structures. A distinction can be made between 'fixed' and 'free' stress: fixed stress is predictable and can in theory be derived by rules, while free stress is unpredictable and has to be lexically listed, also referred to as lexical or phonemic stress. This means that the more regular the stress pattern is, the more predictable it is, while when stress is highly irregular, it has to be stored in the lexical representation of each word separately. Conversely, intonation, marking syntactic structures and focus, is never determined in the lexical representation of a word and depends on the pragmatic context of a word or, across word boundaries, a phrase. There is also a difference between bounded and unbounded stress systems: in bounded stress systems, there is either a maximal distance (e.g. two or three syllables) between each stress, or the stresses are at a particular distance from a boundary. In an unbounded system, stress falls at an unlimited distance of a boundary or another stress. For a theoretical approach on the acquisition of bounded stress systems, see Dresher & Kaye (1990), Dresher (1999) and Tesar & Smolensky (2000).

According to metrical grid theory (Liberman & Prince, 1977; Prince, 1983; Nespor & Vogel, 1989) stressed and unstressed positions alternate at different levels of the prosodic hierarchy as represented in (2.1). The first (lowest) level of the grid represents the syllable level, that is, each syllable is associated with a grid mark. On the second level, each secondary word stress is associated with a grid mark; on the third level, each primary word stress is associated with a grid mark; and on the fourth level, a grid mark is associated with the position of phrasal stress. In the grid, boundaries can be marked for different constituents, ordered from large to small (2.2). The prosodic word is the largest constituent of the three, followed by the metrical foot, as a prosodic word can contain multiple feet; the smallest of the three is the syllable as two syllables can together form a foot.

```
(2.1)                    *                   Phrase (strongest phrasal stress)
                         *                   Prosodic word (primary stress)
         *               *                   Foot (secondary stress)
         *       *       *       *           Syllable
```

```
(2.2)    Prosodic word (PrWd)
                 |
            Foot (Ft)
                 |
           Syllable (σ)
```

2.1.2 Iambic-Trochaic Law

The Iambic-Trochaic Law (ITL: Woodrow, 1909; 1911; 1951) defines how phonetic properties of the speech signal are related to the way in which humans parse this signal into perceptual units. The ITL states that (i) elements contrasting in pitch or intensity naturally form groupings with initial prominence, a trochee or SW-unit and (ii) elements contrasting in duration naturally form groupings with final prominence, an iamb or WS-unit (Hayes, 1985; 1987). This law is based on the principle that the metrical foot has at least one, and maximally one, strong syllable. Quantity sensitivity (QS) can cause skewedness or asymmetries in the occurrence of SW and WS patterns. QS refers to whether heavy syllables, i.e. syllables which are relatively high on a scale of weight counting phonological units, may or may not occur in the weak position of a foot. Hayes (1995) uses a limited inventory of three basic (bounded) foot types: the syllabic trochee, the moraic trochee and the iamb. In the syllabic and moraic trochees, the strong syllable is left edge (SW) while in the iamb, it is right edge (WS).

The question whether the ITL and these asymmetries are 'innate' or language-dependent can be approached experimentally. This can be done by separately manipulating pitch or intensity and duration in experiments testing adults who have different language backgrounds (a.o. Hay & Diehl, 2007). In such a study, Bhatara et al. (2013) found that adult speakers of French and German both abide by the ITL by demonstrating trochaic groupings when intensity or pitch was manipulated, and iambic groupings when duration was manipulated (for infants, see Abboub et al., 2016). However, both effects were found to be larger and more consistent in the German speakers than in the French speakers. This language-specific difference may reflect the function of stress in each language. German has an SW stress pattern at the word level, while French is commonly assumed to have a WS stress pattern not at the word level, but at the phrase level (Grammont, 1965; Dupoux, Pallier, Sebastián & Mehler, 1997; Jun & Fougeron, 2000).

2.1.3 Universal Trochaic Bias hypothesis

The Universal Trochaic Bias (UTB) hypothesis states that all humans show a bias for trochaic units, or SW stress patterns, in their linguistic processing during the early stages of language acquisition (Allen & Hawkins, 1978; 1980). In Allen & Hawkins (1978, but also in Smith, 1973 and Ingram, 1974), the observation was made that children learning English, which has a dominant SW (trochaic) word stress pattern, delete unstressed syllables in their early word productions. Children do this in two specific phonological contexts: (i) if a weak syllable is in word-initial position, e.g. ['pokio] for 'Pinocchio', and (ii) if a weak syllable is adjacent to another light syllable, e.g. [ˈopɪŋ] for 'opening'. Both conditions support the existence of an initial trochaic bias, as these deletions result in a trochaic form: either by ensuring the initial, and sometimes only, syllable is stressed or by truncating a polysyllabic word to a two syllable form of which the first one is stressed.

In a subsequent experimental study, Allen & Hawkins (1980) found that children repeated disyllabic non-words with an iambic pattern as words with a trochaic pattern in the majority of the cases. They additionally conducted a discrimination task in which the children were asked to tell the difference between non-words differing in rhythmic form: iambic or trochaic. Even though the children preferred to produce trochaic words, in perception, they did not have a problem discriminating between these two stress patterns. Other early studies examining the relation between perception and production comprise the work by Echols (1987; 1988). She investigated whether perceptual biases, based on the salience of stressed and final syllables in the linguistic input children receive, can account for the production patterns obtained in two reproduction tasks: a direct repetition task and a learning task. Her data suggest that this is indeed the case (and she argues they support so-called 'autosegmental' phonology), but they can just as well be accounted for by the UTB hypothesis, which can explain the obtained production patterns independently of the linguistic input.

Fikkert (1994) performed a larger study on the truncation patterns in the early language production of children learning Dutch, also a trochaic language. Her more detailed analysis on a larger sample of child speech confirms the observations by Allen & Hawkins (1978), but also adds to it: she demonstrates that in early stages, children do not only delete unstressed syllables but they also adapt words which do not fit the trochaic template in several different ways in order to make them fit this template. She discusses four consecutive stages in development, attested by the modified production of iambic targets: (i) deletion of the initial unstressed syllable, (ii) stress shift from the final to the initial syllable, (iii) production of a disyllabic word with level stress and (iv) production of the iambic target. On the other hand, trochaic targets are hardly ever adapted: in early stages, they are reduced to monosyllables by means of syllable deletion, but they are never adapted by stress shift.

However, Fikkert (1994) points out that the data, although providing evidence for a trochaic bias in English- and Dutch-learning children, do not support a ‘universal’ trochaic bias, as English and Dutch also happen to be trochaic languages. In order to better judge this issue, it would be helpful to have comparison material from one or more non-trochaic languages so as to determine whether an initial trochaic bias is also attested in the early acquisition of children acquiring another type of language than a trochaic language. Such a study was reported on more recently by Adam & Bat-El (2008; 2009), investigating Hebrew which the literature considers to be a predominantly iambic stress language: 75% of Hebrew nouns and 95% of Hebrew verb stems are iambic (although Adam & Bat-El, 2008; 2009 are more careful in concluding Hebrew is clear SW-language, due to stress variation). Their results show that up to approximately the first 100 words, Hebrew-learning children also adhere to the trochaic foot in their early productions. Only later in development do Hebrew-learning children switch to producing forms with the language’s predominant iambic pattern. The authors view this as evidence for the existence of a universal trochaic bias. It must be noted, however, that the findings are based on a small sample of only a few children.

Looking for further evidence from languages with final word stress, given which ones fill this typological slot, the clearest example is constituted by Turkish (Goedemans & van der Hulst, 2009). Topbaş (1997) studied the phonological acquisition by Turkish-learning children and investigated, in particular, their syllable simplification patterns. One of the observed patterns was syllable deletion, but this phonological process was only attested in five out of 22 children. Interestingly, syllable deletion thus seems to be less widespread in Turkish than in other languages for which this pattern has been described, such as for English and Dutch. When looking at Topaş’ Turkish data from a qualitative perspective, it is striking to see that although most disyllabic and trisyllabic words are not truncated, other segmental phonological errors are almost always made in non-final syllables. The final, by default, stressed syllable is correctly produced in most examples given, indicating perceptual salience due to stress plays an important role in learning.

Kopkallı-Yavuz & Topbaş (2000) provide a more detailed analysis of 19 Turkish children’s preferences for syllabic structures in the first 70-100 words and between the ages of 1;3 and 2;0 years. These data seem to confirm the relative lack of typical truncation patterns in Turkish-learning children’s productions, as the most frequent syllabic structure found was the disyllabic word (57% of all productions, against only 19% monosyllabic, 19% trisyllabic and even 5% polysyllabic words). This is in contrast with the widely accepted, and in other languages attested, generalization that young children prefer monosyllables in their early productions. Topbaş (2006) also makes a cross-linguistic comparison of error patterns in children with a phonological disorder and, again, finds that syllable deletion is less frequent in Turkish-learning children than it is, for example, in English- and Spanish-learning children with a phonological disorder. Weak, or unstressed, syllable deletion is a much more frequent error pattern in these children's speech, than syllable deletion in general is in Turkish.

In sum, there appears to be no conclusive evidence yet in favor of or against the UTB hypothesis from the production studies discussed above. Languages with a dominant trochaic pattern, such as English and Dutch indeed show a trochaic bias in early word productions and even some studies on languages with a dominant non-trochaic pattern, such as Hebrew, demonstrate an initial trochaic bias. However, evidence from production studies of Turkish, a non-trochaic language, indicates that the deletion of unstressed syllables in order to obtain a trochaic pattern in early word productions is not wide-spread cross-linguistically. Additional evidence disfavoring the UTB hypothesis comes from a case study by Rose & Champdoizeau (2008) examining the early word productions of a bilingual English-French child not showing any support for the existence of a trochaic bias. They therefore propose to breathe new life into the Neutral Start Hypothesis (e.g. Klein, 1984; Hochberg, 1988) denying the existence of an innate rhythmic bias in language acquisition.

The studies investigating the UTB hypothesis described thus far are, with one exception (Allen & Hawkins, 1980), all about the early productions of children. It is generally assumed in language acquisition research that simplified productions are the consequence of yet to be fully developed perception abilities or of pre-adult grammars (reflecting perhaps universal 'mental' biases or incorrect preliminary generalizations) acting on 'underlying forms' (Gnanadesikan, 1995; Smith, 1973; Levelt, 2012). It is thus of great importance to also test the UTB hypothesis in perception experiments. Moreover, perceptual development can be studied much earlier than productive development, which is important when testing innate biases in acquisition. Furthermore, the UTB hypothesis should be tested on a wider range of both trochaic and non-trochaic languages than thus far has been the case.

2.1.4 Rhythm Class Hypothesis

The Rhythm Class Hypothesis (RCH) is based on the notion of isochrony, i.e. the idea that particular phonetically or phonologically different chunks of speech have similar durations. Languages can be placed into categories of isochronism, with stress-timed (e.g. Dutch), syllable-timed (e.g. Turkish) and mora-timed languages (e.g. Japanese) as the main rhythm classes, described initially by Pike (1945) and Abercrombie (1967). Syllable-timed languages were believed to be characterized by isochrony between syllables, whereas in stress-timed languages interstress intervals are isochronous, and in mora-timed languages, the requirement of isochrony held at the level of the mora. However, this strict three-way version of the RCH has been the topic of debate ever since its formulation and more recently, the dominant view is that there are either more than three rhythm classes (Ramus, Dupoux & Mehler, 2003) or there exists a continuum of rhythmicity (Grabe & Low, 2002).

One important point of criticism of the RCH came from researchers unable to find much phonetic evidence for durational differences between interstress intervals when comparing languages from the three traditional rhythm classes (Dauer, 1983; den Os, 1988). Dauer (1983) proposed that other phonological properties, such as the amount of variance in syllable types (CV, CVC, CCVC, etc.) and the use of vowel reduction do, in fact, distinguish syllable-timed languages from stress-timed languages. It would be these phonological properties which create the perception of different rhythms. According to Ramus, Nespor & Mehler (1999), a continuum of rhythmicity is established by the way in which vowels alternate with consonants and can be described by two durational measures: namely %V (proportion of vocalic intervals) and ΔC (standard deviation of consonantal intervals). The distribution of languages which arises from these measurements is depicted in Figure 2.1 (adapted from Ramus et al., 1999). This distribution shows that Dutch is indeed on the stress-timed side of the continuum with a lower %V, while Turkish is on the syllable-timed side of the continuum with a higher %V, although the languages do not differ much with respect to the ΔC measure. Testing the reality of the RCH has also played a central role in early language acquisition studies and these studies will be reported on in Section 2.2, reviewing the literature on the early sensitivity of infants to rhythm classes.

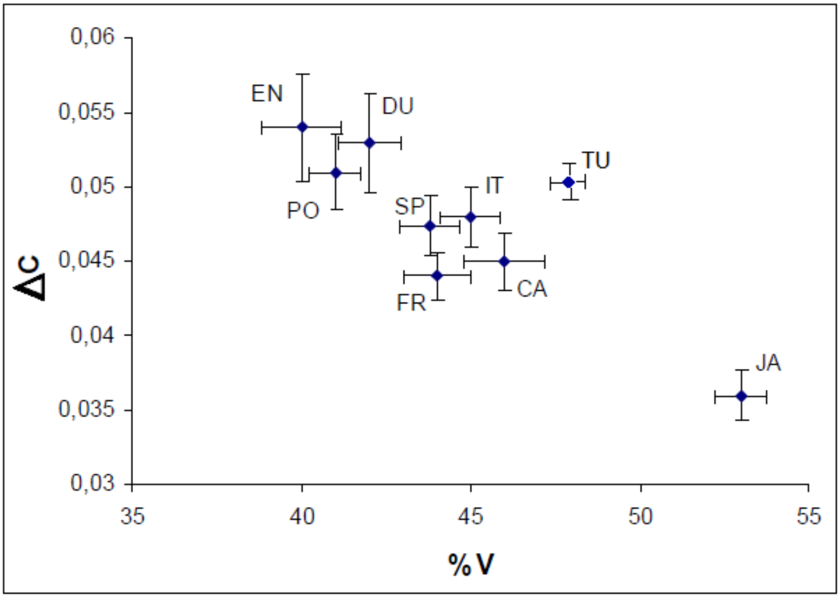


Figure 2.1: *The distribution of languages as measured by %V (proportion of vocalic intervals) and ΔC (standard deviation of consonantal intervals), adapted from Ramus et al. (1999).*

2.2 Rhythmic discrimination and preferences

The focus of this dissertation is on perceptual development and, therefore, the literature on the perception and processing of rhythmic patterns will be reviewed in the next sections: first, the literature on rhythmic discrimination studies in Section 2.2.1 and then the literature on rhythmic preference studies in Section 2.2.2, discussed separately, because they are substantially different. That is, they attempt to answer questions which are intrinsically different. Rhythmic discrimination studies mostly test the lower-level acoustic discrimination of (languages with) different rhythmic patterns which need not be based on language-specific knowledge, whereas rhythmic preference studies aim to test the higher-level phonological processing of different rhythmic patterns, thereby truly addressing issues of native language acquisition. Both are useful for the current study, however, because acoustic processing hints at the role of universal biases in processing, while phonological processing touches on the use of language-specific representations in processing.

2.2.1 Rhythmic discrimination

One of the first investigations into the perception of rhythmic information in early language acquisition was a study by Jusczyk & Thompson (1978). Their results, obtained by using the High Amplitude Sucking (HAS) procedure, suggest that English-learning infants are already sensitive to different stress patterns at 2 months of age, when presented with multiple disyllabic non-words differing in their stress pattern, SW versus WS, trochaic versus iambic. Ten years later, Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison (1988) elaborated on these results by investigating not only English-learning 2-month-old infants, but also French-learning 2-month-olds, and 4-day-old English-learning and French-learning newborns, also using the HAS procedure. Their results show that both 2-month-olds and 4-day-olds can discriminate their native language from a foreign language when presented with low-pass filtered stimuli, in which suprasegmental information is preserved but segmental information is not, coming from full-sentence speech samples. These findings demonstrate that very young infants are able to discriminate two languages based on suprasegmental information only, as opposed to requiring both suprasegmental and segmental information.

Nazzi, Jusczyk & Johnson (2000) then used a different methodology suited for older infants, the Head-turn Preference Procedure (HPP), to test whether English-learning 5-month-old infants could only discriminate languages belonging to different rhythm classes or whether they could also discriminate languages belonging to the same rhythm class. As mentioned earlier, the traditional distinction of rhythm classes is a three-way division between languages with a stress-timed rhythm, such as English and German, with a syllable-timed rhythm, such as Spanish and Italian, and mora-timed languages, for example, Japanese. They found that English-learning 5-month-olds could discriminate languages belonging to different rhythm classes (English or Italian versus Japanese), but that they could not discriminate two foreign languages within the same rhythm class (Italian versus Spanish or German versus Dutch). However, they could discriminate languages within the same rhythm class when one of the languages was their native language. Nazzi and colleagues (2000) suggest that the ability to discriminate languages is thus dependent on the experience infants have with their native language (see also Bosch & Sebastián-Gallés, 1997) and that these results do not necessarily confirm the RCH in infants, because discrimination within a rhythm class is possible as long as one of the languages with that class is familiar to the infants. It should be noted that the full-sentence stimuli in this study were not low-pass filtered, hence, we cannot distinguish whether infants used both segmental and suprasegmental information or whether they can only rely on prosodic information to discriminate languages. It could be the case that the infants used the segmental information to recognize their native language, which would make these particular results less informative about early prosodic development only.

An important question that can be asked about studies in this area is whether the discrimination of different stress patterns by infants is based purely on acoustic information or on higher-level phonological representations. If the former, discrimination would be based on the degree of acoustic salience between stress patterns or languages; if the latter, discrimination would be based on the acquisition of language-specific representations. This issue was investigated by Skoruppa et al. (2009). Because Spanish has word level stress while French only has phrase level 'accent' (Grammont, 1965; Dupoux et al., 1997; Jun & Fougeron, 2000), Skoruppa and colleagues presented 9-month-old Spanish-learning and French-learning infants with either segmentally varied disyllabic non-words or a single disyllabic non-word, which were either stressed on the first syllable (SW) or on the second syllable (WS). The results indicate that Spanish-learning infants could discriminate the SW stimuli from the WS stimuli when presented with segmentally varied non-words, but the French-learning infants could not. The French-learning 9-month-olds could, however, discriminate the different stress patterns when presented with a single non-word. Skoruppa and colleagues interpreted these results as an indication of the difference between processing stress at the phonological level by Spanish-learning infants and processing stress at a purely acoustic level by French-learning infants. This suggests, again importantly for the current study, that a phonetic-phonological property of French underlies these findings. Spanish is a language that requires the development of an abstract representation of stress at the word level in acquisition, while French may well be one which does not require the development of such a phonological representation.

The authors being aware of this latter issue, ran a follow-up experiment with Spanish-learning and French-learning 6-month-olds, i.e. one in which the infants were of a younger age (Skoruppa et al., 2013). The results of this experiment show that at this younger age neither Spanish- nor French-learning infants can discriminate the different stress patterns when presented with segmentally varied non-words, even though both language groups can discriminate the stress patterns when there is no segmental variation present in the stimuli. The authors conclude that while all infants can discriminate different stress patterns acoustically, only infants learning a language in which stress is relevant on a phonological level also learn to represent stress on this higher level of processing. Therefore, only more advanced learners, i.e. older infants, of a language with word stress are able to normalize over segmental variation. I will return to the matter of using single versus multiple words as stimuli in rhythmic discrimination and preference studies in Section 2.3, where I will discuss a number of methodological issues in these type of studies.

2.2.2 Rhythmic preferences

As it was established that newborns are already able to discriminate languages and non-word stimuli based on suprasegmental information, the emphasis of studies on early rhythmic sensitivity shifted from discrimination procedures to preference procedures. In other words, the question was raised whether infants could not only perceive the difference between different stress patterns, but also demonstrate a preference for the stress pattern of their native language. Jusczyk et al. (1993a) explored this question by testing 6- and 9-month-old English-learning infants on their listening preferences using the HPP. They found that infants listened longer to different low-frequent English SW-words than to different low-frequent English WS-words at 9 months of age, but not yet at 6 months. They inferred from this that a language-specific rhythmic preference develops between 6 and 9 months of age. Moreover, at 9 months of age this effect was maintained when presenting infants with the low-pass filtered version of the stimuli, demonstrating that the SW preference at 9 months can also be elicited by presenting the prosodic features of the stimuli only.

As a follow-up experiment to this within-language (English only) study, Jusczyk et al. (1993b) conducted a cross-linguistic study with 6- and 9-month-old English-learning and Dutch-learning infants in which they tested language-specific preferences using the HPP. They first presented both language groups with English and Dutch stimuli, both stress-timed trochaic languages, and subsequently decided to present only the English-learning infants with English and Norwegian stimuli as Norwegian differs from English prosodically. In order to test whether English- and Dutch-learning infants would also base their preference on the segmental properties of the language, half of the number of stimuli contained phonemes or phoneme sequences not occurring or being illegal in the other language. The stimuli were spoken by a balanced bilingual English-Dutch female speaker. The results showed that both English-learning and Dutch-learning infants, but again only the 9-month-old groups, demonstrate a preference for stimuli that match their native language. However, in this study the native language preference was not maintained when presenting English-learning 9-month-old infants with the low-pass filtered version of the stimuli, suggesting that the language-specific preference was mostly based on the segmental properties of the stimuli, as opposed to the prosodic properties which are very similar in English and Dutch.

In the second part of the study, they tested the younger English-learning infants on low-pass filtered stimuli from languages which differ more in their prosodic structure than English and Dutch do, namely English and Norwegian. Norwegian is a stress-timed language similar to English, but different from English as it uses lexical tones. This means that the Norwegian stimuli had a pitch rise on the last syllable of the word more often than the low-pass filtered English stimuli, making them more distinct (Haugen & Joos, 1972). As a consequence, the English-learning infants in this experiment already showed a preference for the stimuli of their native language at 6 months of age.

Herold, Höhle, Walch, Weber & Obladen (2008) investigated rhythmic preferences in infants learning German, a language with trochaic word stress, also by using the HPP. They compared a group of preterm (mean gestational age [GA] of 28 weeks) very low birth weight (<1500 g) German-learning infants with a control group of full term (mean GA of 40 weeks) German-learning infants at the (corrected) ages of 4 and 6 months on their preference for different stress patterns. Although the preterm infants did not show any sign of the expected preference for listening to the SW version of the disyllabic non-word /'gaba/ over the WS version /ga'ba/, at the corrected ages of 4 and 6 months, the full term infants did demonstrate a preference for the SW condition over the WS condition at both 4 and 6 months of age. However, the fact that the 4-month-old infants in this study were exposed to a familiarization phase consisting of only SW stimuli prior to testing, makes it difficult to conclude that the 4-month-old German-learning infants truly showed a preference for the SW condition based on their outside-the-lab-experience with the stress pattern of their native language. The older group of 6-month-old German-learning infants in this study were, however, not exposed to a familiarization phase prior to testing, hence their longer orientation times (OTs) during the SW condition can be interpreted as a preference for the stress pattern of their native language at this age. The upshot seems to be that German-learning infants do not show a preference for their native language based on a native phonological representation yet at 4 months of age, but they do at 6 months of age.

Dutch-learning infants have only been tested at 9 months of age thus far (Jusczyk et al., 1993b) and only using stimuli with segmental variation in an experiment in which the manipulated variable was not stress, but the non-stress phonetic and phonotactic patterns. This is of course exactly where the motivation for this dissertation lies. It is not just concerned with acquisition, but also specifically with providing data from Dutch-learning infants in the critical age range described in the literature, namely between 4 and 8 months of age. As all of the languages for which a native SW preference has been found are also languages with a predominant SW stress pattern, it will still be difficult to disentangle the Dutch results from a possible confirmation of the UTB hypothesis, which states that there is an universal perceptual bias towards SW units over WS units in all human infants (Allen & Hawkins, 1978; 1980). Therefore, a non-trochaic language has to be added to the design of the current study.

Thus far, the only non-trochaic languages discussed elaborately in the infant perception studies were Spanish and French (for French bilingual infants, see Abboub et al., 2015; Bijeljac-Babic et al., 2012; 2016). Spanish-learning infants have mostly been tested in discrimination procedures (Skoruppa et al., 2009; 2013,) instead of in preference procedures. Moreover, Spanish does not really have a predominant SW or WS stress pattern: within the category of disyllabic words, there are more trochaic than iambic words, but not very impressively so, the distribution being 60% versus 40% respectively (Pons & Bosch, 2010). We therefore have to conclude that in a preference experiment, Spanish is not the right language to test the reality of the UTB hypothesis on. French, however, does have a dominant WS pattern, but this pattern seems to manifests itself at the phrase level, not at the word level (Grammont, 1965; Dupoux et al., 1997; Jun & Fougeron, 2000). French also does not seem the right language to investigate when attempting to gather evidence in support of or against the UTB hypothesis.

Among the languages which the literature reports have a dominant WS pattern at the word level, Turkish has default stress on the last syllable of the word, which is either the final syllable of a non-morphologically complex word or, more often, a suffix due to the agglutinative character of the Turkish language. Turkish does, however, have exceptions to this final stress pattern, mostly in proper names and under certain morphological conditions, such as pre-stressing suffixes (e.g. Sezer, 1983; Inkelas, 1999; Kabak & Vogel, 2001; Inkelas & Orgun, 2003; Göksel & Kerslake, 2005). If, indeed, Turkish can be considered to be a language with word-final stress, and issue that will be further addressed in Chapter 2, this implies it can be used as an adequate test case for the UTB hypothesis, filling the typological gap discussed above. Therefore, a study of Turkish has been included in this dissertation: it has the aim of finding out whether Turkish-learning infants between 4 and 8 months of age show a WS preference based on the dominant word-final stress pattern of their native language, show no bias at all, or alternatively show a universal trochaic bias.

I will first, however, elaborate on a number of methodological issues that can be raised for the above rhythmic discrimination and preference studies in Section 2.3 and provide a review of the literature on speech segmentation by infants in Section 2.4, before proceeding to the description of the phonological properties of the languages under investigation in this dissertation, Dutch and Turkish, in Chapter 3.

2.3 Methodological issues

From the literature review so far on rhythmic discrimination and preference studies we can conclude that studies on infants learning Germanic languages, predominantly German and English, and Romance languages, primarily Spanish and French, have provided rather mixed results. For German-learning infants, a true native SW preference was not found at 4 months of age yet, but it was found at 6 months of age (Herold et al., 2008). For English-learning infants a native SW preference was only found under specific conditions at 6 months of age, but it was found under several conditions at 9 months of age (Jusczyk et al., 1993a; 1993b). Spanish-learning infants were found to be able to discriminate different stress patterns in all conditions at 9 months of age (Skoruppa et al., 2009; 2013), whereas French-learning infants could only discriminate the different patterns at 6 months under certain conditions (Skoruppa et al., 2013), but not at 9 months under more challenging conditions (Skoruppa et al., 2009).

It is the purpose of this section to discuss these results in light of the methodological properties of and differences between the studies. The first issue I will discuss is the matter of using single versus multiple word stimuli in Section 2.3.1, then I will discuss the issue of native versus non-native stimuli in Section 2.3.2 and, subsequently, the issue of the acoustic manipulations performed on the stimuli in the different studies will be discussed in Section 2.3.3. Here and in general, a relatively large part of this dissertation will be focused on the importance of methodological issues, not only regarding the nature of the stimuli, but also regarding the collection and interpretation of infant data gathered in this type of attentional preference based experiments.

2.3.1 Single versus multiple word stimuli

The studies discussed differ in the type of stimuli with which infants are presented: some use multiple tokens of a single non-word (Herold et al., 2008; Skoruppa et al., 2013), whereas others (Skoruppa et al. 2009; Jusczyk et al., 1993a; 1993b) use multiple (non-)words. The studies that use a single non-word find discrimination of, and preferences for, (native) stress patterns as young as 6 months of age. However, the other studies fail to find discrimination or native preferences up to 9 months of age. In both studies by Jusczyk et al. (1993a; 1993b) infants were presented with multiple low-frequent words containing segmental variation (except for the condition in which infants listened to low-pass filtered stimuli). This segmental variation may explain the lack of a native preference at 6 months of age in the English-learning infants, an interpretation which is compatible with the study by Skoruppa et al. (2013) in which 6-month-old Spanish-learning infants are unable to generalize stress patterns over segmentally varied stimuli. Jusczyk et al. (1993a) did test 9-month-old infants with a low-pass filtered version of the stimuli, hence without segmental variation, as well, but unfortunately not at 6 months of age. In other words, we do not know whether 6-month-old English-learning infants would show a native rhythmic preference in the absence of segmental variation, but my speculation would be that they will.

Some support for this speculation can be found in the cross-linguistic experiment by Jusczyk et al. (1993b) in which English-learning infants at 6 months of age, provided that the languages presented are sufficiently distinct prosodically, such as English and Norwegian, do show a native preference when presented with low-pass filtered stimuli. This suggests that 6-month-old infants still have difficulties processing suprasegmental information in the face of segmental variation, but not when this variation is absent. Unfortunately, Jusczyk and colleagues (1993b), although addressing both English and Dutch, did not manage to test Dutch-learning infants on the same task, hence it is unknown whether Dutch-learning 6-month-olds would display similar results as the English-learning infants in this last type of experiment. Nonetheless, all of these studies show that for infants up to 6 months of age stimuli without segmental variation are more suitable for eliciting native rhythmic preferences.

2.3.2 Native versus non-native stimuli

Another issue is that of the speaker delivering the stimuli. The studies by Mehler et al. (1988) and Nazzi et al. (2000) clearly demonstrate that including stimuli from the native language, i.e. stimuli spoken by a speaker of the native language of the infants, can be crucial for discrimination. This notion was unfortunately not taken into account in the later discrimination studies by Skoruppa and colleagues (2009; 2013), in which different language groups of Spanish- and French-learning infants were presented with stimuli recorded by a speaker of Spanish, the native language of only one of the infant groups. The reported reason for this is that a native speaker of French would not have been able to produce the different stress patterns properly, supposedly because of so-called stress 'deafness'. This is presumed to be the lowered ability to lexically encode stress in speakers of languages with either fixed word or phrase level stress (Dupoux, Peperkamp & Sebastián-Gallés, 2001), the underlying way of thinking being that if stress is not phonemic at the word level, it is not stored in the lexical representation of words.

To find further support for their decision to present the infants with Spanish-spoken stimuli, Skoruppa et al. (2009 ; 2013) had the Spanish-spoken stimuli mimicked by a native speaker of French and both the Spanish-spoken and French-spoken stimuli were then presented to adult speakers of French in a perception task similar to that of Dupoux and colleagues (2001). The results showed that the speaker did not affect their performance on the task. Skoruppa and colleagues (2009) conclude that the French-learning infant's lack of discrimination at 9 months of age cannot be caused by the fact that the stimuli were spoken by a speaker of a (to them) foreign language. However, the processing of non-native stimuli by adults is not necessarily representative for how infants process this type of stimuli, and a confounding effect of non-native stimuli in infant experiments can therefore not be ruled out.

In both studies by Skoruppa and colleagues (2009; 2013), there was arguably an advantage for the Spanish-learning infants processing native stimuli as compared to the French-learning infants processing non-native materials. As discrimination of the stimuli is intrinsically a prerequisite for preference studies, this inequality would preferably be ironed out in future cross-linguistic rhythmic preference studies. This can be accomplished in one of two ways: either both language groups should be presented with stimuli spoken by a speaker of a (to them) foreign language, in other words, a speaker of a third language which is non-native for both language groups, so as to at least provide an equal starting point for both language groups. Another approach would be to present each language group with stimuli spoken by a speaker of their native language, thereby not providing one of the groups with an advantage or disadvantage. However, this last option might lead to another confound in cross-linguistic studies, namely the fact that the stimuli presented to the different language groups cannot be completely identical, as this option may involve different speakers for the different sets of stimuli. This problem can be solved by having a balanced bilingual speaker of both languages record the stimuli, as was done by Jusczyk and colleagues (1993b).

2.3.3 Acoustic manipulation of the stimuli

Related to the previous issue of native versus non-native stimuli is the matter of the acoustic manipulations performed on the stimuli presented to the infants. In all of the studies discussed above, the goal was to present infants with stimuli differing in word stress patterns: trochaic (SW) versus iambic (WS). However, the way in which this stress distinction was realized phonetically differs substantially between the studies. In some cases the acoustic cues present in the stimuli are indeed related to the phonetics of the language of the speaker by whom the stimuli were recorded, as is the case in the studies by Skoruppa et al. (2009; 2013). In their Spanish-spoken stimuli, all three dimensions of stress: pitch, duration and intensity were manipulated by the speaker resulting in SW and WS stimuli, of which the stressed syllables were always higher in pitch, longer in duration and higher in intensity than the unstressed syllables. Whether all of these acoustic cues are also used in French when producing phrasal 'accent' (which is realized by final lengthening without an increase in pitch or duration according to Vaissière (1991) and Dupoux, Sebastián-Gallés, Navarrete & Peperkamp (2008)) is left undiscussed by the authors, but the rationale seems to be that if all possible acoustic cues to stress are present in the stimuli, French-learning infants could in principle pick-and-choose one or multiple cues for their native perception of prominence.

Other studies, such as Jusczyk et al. (1993b) and Herold et al. (2008) do not provide any acoustic measures of the stimuli whatsoever. The former simply state that for the languages under investigation, English and Dutch, literature shows that they display similar pitch, duration and intensity correlates of stress. Precisely this assertion, however, is not reflected in acoustic measures of stress in both languages. Although this is less harmful in this particular study using a balanced bilingual speaker for recording the stimuli (and, with that, assuming the relevant acoustic cues to stress will be present in the stimuli for both English- and Dutch-learning infants), providing detailed acoustic measures of the stimuli used in rhythmic discrimination and preference studies still is highly preferable. Throughout the rest of this dissertation, I will therefore always be alert to and critical of the acoustic manipulations performed on the stimuli used in my own work as well as in the work of others. For an overview of the methodological properties of the studies discussed thus far, see Table 2.1. Moreover, all of the methodological issues above also hold for the speech segmentation literature which will be discussed in the last section of this chapter, Section 2.4.

erview of the rhythmic discrimination and preference studies discussed thus far (more studies follow in Chapte
N), French (FR), Italian (IT), Spanish (SP), German (GE), Dutch (DU), Japanese (JA) and Norwegian (NO).

)	*Language(s)*	*Method*	*Stimuli*	*Speaker*	*Acoustic manipulatior*
ompson (1978)	EN	HAS	multiple non-words	EN	pitch, duration, intens
1988)	EN, FR	HAS	full sentences	EN, FR	low-pass filtered
000)	EN	HPP	full sentences	IT, SP, GE, DU, JA	no manipulation
l. (2009)	FR, SP	HPP	single/ multiple non-words	SP	pitch, duration, intens
l. (2013)	FR, SP	HPP	single/ multiple non-words	SP	pitch, duration, intens
1993a)	EN	HPP	multiple words	EN	low-pass filtered
1993b)	EN, DU	HPP	multiple words	EN, DU, NO	low-pass filtered
2008)	GE	HPP	single non-word	GE	not reported

2.4 Speech segmentation

For an adult native speaker of a certain language, it must be a fairly easy task to break up the speech stream into phrases and words when listening to that language. In fact, adult speakers hardly ever seem to be aware that the speech they listen to every day is physically only a continuous stream of sound instead of a concatenation of well-demarcated words. One of the reasons for this unawareness is that adult speakers possess a lexicon, a mental dictionary, in which a vast amount of the words of their language are stored: not only their form, the phonology of the word, but also their meaning. Adults, therefore, can make use of a strategy called top-down processing when listening to their native language. Completely different for these same adults, of course, is the task of listening to a foreign, unknown language. Their native lexicon is of not much use when processing a language with a completely different dictionary content and they now have to resort to a different strategy, namely bottom-up processing. The first challenge in this is not to find out the meaning of words, but first and foremost the word forms have to be sorted out: where do words begin and where do they end?

Listening to a foreign language makes adult speakers aware of the fact that speech does indeed not consist of concatenated words, but of a continuous, very fast, stream of sounds. For newborns, the experience is supposedly very similar. They do not possess a lexicon yet and, thus, until that moment, are also bottom-up processors trying to figure out how to break up the speech stream into smaller chunks representing, first, phrases and, later, words. Only after discovering the word forms can they start mapping these hypothesized words onto meanings and, with that, start building up a lexicon to become the efficient top-down processors adult language users are. In Section 2.4.1, I will review the literature on the development of this ability called word segmentation and which is aided by different cues, before proceeding to discuss the Metrical Segmentation Hypothesis (MSH) for word segmentation in Section 2.4.2.

2.4.1 Word segmentation by infants

The reason why it is not self-evident to segment the speech of a foreign, unknown language is because cross-linguistically, there is not one unique cue which marks word boundaries (Saffran et al., 1996). Even so, across languages, there are multiple cues that may aid young learners to find such boundaries. Among these are statistical information, the properties of functional (versus content) morphemes, phonotactics and prosody. Some of these cues are categorical in some languages, e.g. phonotactics in Dutch, but most of them are probabilistic, i.e. none of them can predict word boundaries completely correctly on its own. It therefore has become apparent that the integration of different cues is a prerequisite for accurate word segmentation to take place. The different cues and their necessary integration are discussed below.

Statistical learning

Saffran et al. (1996) have demonstrated that independently of other, acoustic or prosodic, cues 8-month-old infants can use the distribution of syllables in an artificial language (AL) to infer word-like units. In these researchers' artificial language, some syllable pairs had a so-called transitional probability (TP) of 1, which means these syllables were always adjacent in the AL, whereas other syllable pairs only had a TP of .33, which means these syllables were only adjacent in one third of the cases. The infants listened to this language during a familiarization phase of 2 minutes after which they were presented with both 'words' (TP of 1 in the familiarization phase) and 'part-words' (TP of .33 in the familiarization phase) in a test phase. The infants showed a (novelty) preference for the part-words during the test phase, which can be interpreted as them having segmented the actual words during familiarization. The authors demonstrated that infants are sensitive to the serial order of syllables, and furthermore, that this effect was independent of the frequency of the syllables in the familiarization phase.

Thiessen, Hill & Saffran (2005) and Johnson & Tyler (2010) qualified these findings. They showed that when the complexity of the statistical learning task is raised by increasing the number of possible words to segment or segmenting words of varying length, younger infants need more acoustic support to be able to segment the words or are unable to segment the words at all. Thiessen and colleagues compared the statistical learning abilities of 7-month-old infants in two conditions: infant directed speech (IDS) versus adult directed speech (ADS). IDS was in this study typically characterized by the use of a larger pitch range compared to ADS. Even though the IDS speech did not provide the infants with specific prosodic cues towards the detection of the words in the AL task, they still only managed to segment the words in this condition and not in the more monotonous ADS condition. The authors conclude that IDS thus facilitates word segmentation for infants. The means by which this effect is instantiated are nevertheless unclear: in natural IDS, speech tends to be slower, phrases tend to be shorter and there is usually more repetition of words; however, none of these characteristics were present in this AL task. Another explanation would be that the use of IDS heightens the infants' attention during the familiarization phase, which in turn may facilitate learning in general.

Functional morphemes

Although statistical learning seems to be a powerful mechanism, only taking statistics over the frequency of co-occurring syllables may lead to erroneously hypothesized words. For example, determiners in most languages often co-occur with nouns and the set of high frequent nouns in the input to young infants is fairly limited, which causes high-frequent co-occurrences of determiners with particular nouns. If infants only kept track of these co-occurrences, they might chunk the determiner and noun together as one word. However, we know from word segmentation studies that this is not what infants do. Infants track the frequency of occurrence of individual syllables or morphemes, such as determiners, and conclude that if these elements are highly frequent on their own, they are not part of a word but instead signal the boundary of another word, which in turn facilitates word segmentation. This has been demonstrated for both English- and French-learning infants at 8 months of age in studies by Shi, Cutler, Werker & Cruickshank (2006a) and Shi & Lepage (2008), respectively, but similar results are known about German-learning infants' detection of functional morphemes around 8 months of age form work by Höhle & Weissenborn (2003).

The study by Shi et al. (2006a) demonstrates the robustness of the phonological representations of words in 8-month-old infants versus those in 11-month-old infants. For the English-learning 11-month-olds in their experiment, only the determiner 'the' facilitated segmentation, while the fake determiner 'kuh' did not, whereas for the 8-month-olds both 'the' and 'kuh' facilitated word segmentation. According to the authors, this tells us something about the specificity of the phonological representations of words at different stages in development. This is of interest as more detailed segmental phonological representations may also facilitate word segmentation in a different way. Allophonic variation, the phenomenon of phonemes' phonetic realization varying dependent on their position in the word, may also inform word boundaries (Johnson & Jusczyk, 2001). If infants are able to distinguish allophones and have learned about their relationship to the word boundary, this may be an informative source for word segmentation as well. Jusczyk, Hohne & Bauman (1999) have shown that this ability develops between 9 and 10.5 months and that infants at 10.5 months are indeed able to use allophonic information for segmentation. These age-specific results on the use of allophonic variation for word segmentation nicely align with the results from the study by Shi et al. (2006a) on the development of detailed segmental phonological representations between 8 and 11 months of age.

Phonotactics and prosody

Another cue which can be used for word segmentation and also involves statistical learning mechanisms is phonotactics. Phonotactics determine which phoneme sequences are legal and illegal within syllables in a particular language and are hence language-specific. However, infants can use phonotactics earlier than they have learned the exact phonological rules of their native language for word segmentation by calculating bigram frequencies over the phonemes in the input. It is assumed that high frequent bigrams represent within-word sequences, whereas, low-frequent bigrams represent between-word sequences. Friederici & Wessels (1993) found that Dutch-learning 9-month-old infants can indeed use phonotactic cues for word segmentation. Mattys, Jusczyk, Luce & Morgan (1999) confirm this result by presenting results on English-learning 9-month-old infants who show they are sensitive to the difference between well-formed and ill-formed words, which are distinguished by the position of a consonant cluster, which is illegal within a syllable in English.

Mattys and colleagues (1999) investigated the integration of phonotactic cues with prosodic cues in 9-month-old English-learning infants. They found that prosodic cues can heavily influence the segmentation of words containing different phonotactic cues and that prosodic cues are actually leading in word segmentation at this age. Johnson and Jusczyk (2001) found similar results for 8-month-old English-learning infants in an experiment in which they presented the infants with conflicting cues by pitting phonotactic cues against stress cues. They also found that infants rely more on stress cues than on phonotactic cues at 8 months of age. This led to the idea that stress cues might be the first cues used by infants to inform word segmentation.

Testing the latter idea, Thiessen & Saffran (2003) conducted a study with 7- and 9-month-old English-learning infants in which they also presented conflicting phonotactic and stress cues to the infants. Their results show that infants rely on phonotactic cues at 7 months of age and on stress cues at 9 months of age. Conversely, in a later study, Thiessen and Saffran (2007) found that when 7-month-old English-learning infants are trained in a familiarization phase containing informative stress cues, they can and do use stress as a segmentation cue.

In sum, it seems that at least the following obtains: infants rely on a single cue to word segmentation until 9 months of age, although which cue they use seems to be dependent on the exact circumstances of the experiment. From

9 months of age onwards, infants seem to have learned to integrate different cues and then weigh multiple cues in finding the most optimal segmentation of the speech stream into words (Mattys et al., 1999; Thiessen & Saffran, 2003).

The studies discussed so far leave an important question unanswered: can stress cues be used early on (i.e. before 8 months) as a bootstrapping mechanism into speech segmentation, or only later on (i.e. after 8 months) once the initial steps into word segmentation have been made based on other cues, such as statistical learning? There is evidence in favor of the use of rhythmic structure for speech segmentation. This idea is, for example, formulated by Nazzi & Ramus (2003) and proposes that infants' early sensitivity to rhythm allows them to tune into the rhythm type of the language they are hearing and, with that, they can develop the proper word segmentation strategy, i.e. foot-based segmentation for languages on the stress-timed end of the continuum and syllable-based segmentation for languages on the syllable-timed end of continuum. Stress-based languages, however, are thought to be subject to yet a different idea. The hypothesis put forward for these languages is that listeners can use the language-specific stress pattern of their native language, e.g. word initial stress, to inform word segmentation. This is known as the Metrical Segmentation Hypothesis (MSH): it will be discussed more elaborately in the following section.

2.4.2 Metrical Segmentation Hypothesis

The Metrical Segmentation Hypothesis (MSH) was originally put forward by Cutler & Norris (1988) under the assumption that speakers use a parsing strategy based on the metrical representation they have for words in their native language. Where metrical representation here means something in the neighborhood of SW versus WS word-like units. This assumption was supported by data from adults' word segmentation research. It was indeed found that stressed syllables facilitate word segmentation by providing a marker of the word-onset in speakers of both English (Cutler & Norris, 1988) and Dutch (Vroomen & de Gelder, 1995; Quené & Koster, 1998). Not much later, evidence for the use of metrical stress in segmentation was also found in infant studies. Jusczyk, Houston & Newsome (1999) demonstrated that English-learning 7.5-month-olds already use the language's predominant trochaic (SW) pattern for word segmentation. This implies that stress cues can actually inform word segmentation very early on in acquisition (see also Kooijman et al., 2005; 2009; 2013). Additionally, evidence for the hypothesis formulated by Nazzi & Ramus (2003) to the effect that French-learning infants do not use a metrical unit for word segmentation was provided by results demonstrating that the rhythm type of the language which infants receive as input is linked to (in the case of French-learning infants), the absence of rhythmic preferences and, subsequently, to the syllabic-unit French-learning infants use for speech segmentation (Nazzi, Bertoncini & Mehler, 1998; Nazzi, Iakimova, Bertoncini, Frédonie & Alcantara, 2006, see also: Goyet et al., 2010; 2013; Nishibayashi et al., 2015).

Still, as the status of stress in French is being debated, because it only has phrase final prosodic demarcation and there is evidence of French adults being stress 'deaf' (Dupoux et al., 1997), the French data do not form a very strong argument for the language-specificity of the trochaic bias found in English- and German-learning infants. They also do not provide a very strong argument for the language-specificity of the MSH, as all of the studies on languages with a statistically predominant stress-initial pattern allow an alternative interpretation. That is, a stressed syllable might cue a word-onset in all stress languages, regardless of whether the language's dominant stress pattern is SW or WS. In sum, there appears to be no conclusive evidence so far that learners of final stress languages can use the WS stress pattern for word segmentation. It requires data from a cross-linguistic study comparing the acquisition of a word level final stress language, such as Turkish (Kabak & Vogel, 2001; Inkelas & Orgun, 2003), with that of a statistically speaking initial stress language, such as Dutch (Kager, 1989; Trommelen & Zonneveld, 1989), to shed further light on the different hypotheses discussed in this literature review.

Goal

This study aims at testing the language-specificity hypothesis of metrical segmentation for native language acquisition in infants aged 4, 6 and 8 months. It investigates the emergence of rhythmic biases and word segmentation in infants learning metrically opposed languages: Dutch (initial/ prefinal stress) and Turkish (final stress). I will do this by testing rhythmic preferences and metrical segmentation in three experiments, reported on in Chapters 4-7. In the next section I will discuss the general predictions, the research questions and the more specific predictions related to these questions as relevant to the current study.

General predictions

The first general prediction for this study is that infant's language-specific preferences for rhythmic units develop around 6 months of age as a function of the distributional patterns in the input they receive. More specifically, I hypothesize that infants analyze the unsegmented input statistically in terms of rhythmic bigrams: pairs of syllables (strong-weak or weak-strong) and the alignment of these pairs with phrase edges (initial or final). The idea is that infants first represent this distributional knowledge as under- or overrepresented bigrams, before extracting a rhythmic unit or foot (trochaic or iambic) when one of the bigrams is overrepresented. This approach offers a new account of the development of language-specific rhythmic preferences in infants stating that the acquisition of a rhythmic unit is based on continuous speech instead of a 'protolexicon' (e.g. Swingley, 2005), and thereby complementing previously discussed accounts related to rhythm class, such as the proposal by Nazzi & Ramus (2003).

The second general prediction is that this distributionally-based knowledge of rhythmic bigrams also underlies infants' metrical segmentation abilities which are expected to develop around 8 months of age. Once a certain rhythmic unit is extracted, it can subsequently be used to inform word segmentation. That is, a transition within an overrepresented bigram is interpreted as a continuity cue, while a transition within an underrepresented bigram is interpreted as a discontinuity or segmentation cue. Besides these language-specific (dis)continuity cues, there may also be universal (dis)continuity cues at play in word segmentation. Stress clash, i.e. adjacent stressed syllables as discussed earlier in this literature review, may be used as a universal discontinuity cue, whereas the opposite phenomenon of stress lapse, i.e. adjacent unstressed syllables, may universally cue continuity.

To test these predictions I investigate Dutch-learning infants, expected to show a trochaic bias at 6 months of age, similar to German-learning infants (Herold et al., 2008), and Turkish-learning infants, who are expected to show an iambic bias at 6 months of age. In Dutch, SW bigrams are overrepresented (at the left edge) compared to WS bigrams, whereas in Turkish, WS bigrams are overrepresented (at the right edge) compared to SW bigrams (van Ommen, 2016), hence, demonstrating a reverse distribution. If infants of both language groups demonstrated these language-specific preferences at 6 months of age, I expect to find the use of these language-specific stress patterns for word segmentation at 8 months of age. Thus, Dutch-learning infants regard a transition from a strong to a weak syllable (SW, 'trochaic') as a continuity cue, which leads to chunking of SW-units, while a transition from a weak to a strong syllable (WS, 'iambic') is taken as a discontinuity cue, which leads to splitting of WS-units in a word segmentation task. For Turkish-learning infants, the opposite pattern should hold: SW units are split based on the thereby provided discontinuity cue, while WS units are chunked, based on the continuity cue they provide in Turkish.

Research questions and more specific predictions

RQ1: Do Dutch-learning infants and Turkish-learning infants show a language-specific rhythmic preference?

I predict that Dutch-learning and Turkish-learning 4-month-olds fail to show a bias, whereas at 6 months of age both Dutch and Turkish language groups should show a native trochaic and iambic, preference, respectively.

RQ2: Do Dutch-learning infants and Turkish-learning infants use language-specific rhythmic patterns for word segmentation?

If indeed both language groups display a language-specific rhythmic preference at 6 months of age, I predict that Dutch-learning infants regard a strong syllable as the beginning of a word-like unit, a word-onset, whereas Turkish-learning infants regard a strong syllable as the end of a word-like unit, a word-offset, at 8 months of age.

RQ3: Do Dutch-learning infants and Turkish-learning infants use universal cues, such as stress clash and lapse, for word segmentation?

If one or both language groups do not display a language-specific rhythmic preference at 6 months of age, I predict that both Dutch- and Turkish-learning infants do use universal cues, such as stress clash and lapse as discontinuity and continuity cues, respectively, in word segmentation.

Chapter 3
Languages: Dutch and Turkish

3.1 Introduction

This dissertation aims to be a methodologically sound study, taking into account a number of methodological issues which have not always been allowed for in previous studies, on the early acquisition of stress and the use of stress cues for initial word segmentation. In order to be able to execute such an examination, it is both theoretically and practically important to be aware of the phonological and phonetic properties of the languages under investigation. In a cross-linguistic study, the same experimental material is ideally used for both language groups. It is therefore important to describe the phonological and phonetic properties of the languages prior to commencing such a study. This is, on the one hand, necessary for the formulation of exact hypotheses regarding each of the experiments and, on the other hand, for the creation of the stimuli used in the experiments. In Chapter 1, it was pointed out that stress in many languages is an 'edge' phenomenon, where it can be placed at either the left or the right edge of the word, according to language-specific rules. In the current study, the former type of language is presumed to be an SW language, and the latter a WS language. It was also pointed out that thus far in experimental approaches mainly SW languages have been the subject of investigation. This dissertation aims to add data from both an SW language, Dutch, and more importantly, a WS language, Turkish. These languages were also chosen because of the presumable experimental accessibility to infants learning these languages.

Among the available languages for such a study, Dutch and Turkish were chosen as these languages are, rhythmically speaking, the mirror image of one another: Dutch, as an SW language, has predominantly prefinal word stress; however, statistically speaking, this rebounds upon initial stress due to the high incidence of disyllabic words in the language. Schematically, SW is statistically dominant over WS and other patterns. Turkish, on the other hand, has statistically dominant final stress, with WS being dominant over SW; differently positioned stress occurs in circumstances to be described later in this chapter. Moreover, both languages have these systems as default stress, but also allow exceptions. Allowing exceptions can be of importance, as studies on the acquisition of languages without word level stress, such as French, have demonstrated that the absence of lexical variability (distinctiveness) may lead to a minor role for word stress in the acquisition process. In this chapter, the languages' segment inventories (and additionally the phenomenon of vowel harmony in the case of Turkish), morphology and syllable structure, rhythm class, word- and phrase level stress, and the phonetic realization of stress are described, first for Dutch and then for Turkish.

3.2 Dutch

Dutch (Nederlands: Trommelen & Zonneveld, 1989; Booij, 1995; Gussenhoven, 2009) belongs to the West-Germanic language family and has German and English as two of its closest relatives. The language has about 23 million native speakers, the great majority of which resides in two countries: approximately 16 million speakers in the Netherlands and 6 million in Belgium. The language has many variants, of which some are spoken outside the Netherlands and Belgium, in Surinam and the former Dutch Antilles. The two standard variants are Netherlandic Dutch, as spoken in the Randstad area of the Netherlands, and Flemish Dutch as spoken in Flanders, the northern part of Belgium. As all Dutch participants in this study were recruited in Utrecht, which is part of the Randstad area, the properties of standard Netherlandic Dutch are described below.

3.2.1 Phoneme inventory

The consonant inventory of Dutch consists of 18 consonants /p, b, t, d, k, f, v, s, z, x, h, m, n, ŋ, r, ʋ, j, l/ and 16 vowels, of which 13 are monophthongs /a, e, i, o, u, ø, y, ɑ, ɛ, ɪ, ɔ, ʏ, ə/ and three are diphthongs /ɑu, ɛɪ, œy/. The language distinguishes between tense and lax vowels in the form of short and long vowels, respectively, but non-high tense vowels can also be long when stressed. Vowels can also be reduced to schwa when they are unstressed, but this is optional and subject to positional, quantitative and qualitative constraints (Kager, 1989). In standard Netherlandic Dutch some long vowels are produced with diphthongization: /e^{j}, o^{u}, ø^{y}/.

3.2.2 Morphology and syllable structure

Dutch abounds in monosyllabic words and has prefixes as well as suffixes, comprising both derivational and inflectional affixes. The language also allows for elaborate compounding. Regarding syllable structure in non-morphologically complex words, Dutch has many different syllabic templates, as an onset is not obligatory and a coda is allowed (see examples in 3.1). The onset and coda clusters can even amount to three and four consonants, respectively (see examples in 3.2), but the consonants used in clusters of more than two consonants are mostly restricted to /s/ and /t/, which can be attached to the word-onset and -offset as satellites or 'appendices'.

(3.1)	ei	/ɛɪ/	'egg'
	kat	/kɑt/	'cat'
(3.2)	straks	/strɑks/	'later'
	herfst	/hɛrfst/	'autumn'

(Trommelen & Zonneveld, 1989)

3.2.3 Rhythm class

Dutch is categorized as a stress-timed language, similar to English and German. Dauer (1983) states that it demonstrates all features of a typical stress-timed language: it has complex clusters, vowel reduction (although to a lesser extent than English) and variable stress positions. Durational and perceptual studies confirm the classification of Dutch as a stress-timed language together with English and German (Ramus et al., 1999; Grabe & Low, 2002; Ramus et al., 2003; White, Mattys & Wiget, 2012).

3.2.4 Stress

Word stress

When considering polysyllabic words, it turns out that Dutch is a quantity sensitive language, which means the weight of the syllables in a word plays a large role in the assignment of stress in Dutch words. Because of this, it is important to control for the weight of the syllables which are used as stimuli in a study investigating listening preferences for certain stress patterns. Stress is typically placed on the penultimate (prefinal) syllable of a polysyllabic word which, due to the high frequency of disyllabic words, often coincides with the initial syllable depending on the morphological status of the words: poly- or monomorphemic (see examples 3.3 to 3.5, respectively).

(3.3)	verwárming	/vər'ʋɑrmɪŋ/	'heating'
(3.4)	kanárie	/ka'nari/	'canary'
(3.5)	ólie	/'oli/	'oil'

There are, however, exceptions to this pattern, mainly due to quantity sensitivity (QS). When a word is (minimally) trisyllabic and has an open penultimate syllable and a closed final syllable, stress is assigned to the antepenultimate syllable, which - again due to the higher frequency of trisyllabic words over longer words - often is the initial syllable of the word (see example 3.6). Furthermore, when a word has a super heavy syllable, such as VVC (banaan 'banana') or VCC, in word-final position this syllable attracts stress and, hence, this is the only type of word with default final stress in Dutch (see example 3.7: Kager, 1989; van Oostendorp, 1995), besides lexically specified exceptions with stress on non-super heavy syllables in final position.

(3.6)	márathon	/'mɑratɔn/	'marathon'
(3.7)	elegánt	/elə'xɑnt/	'elegant'
(3.8)	formúle	/for'mylə/	'formula'
(3.9)	Bélgië	/'belxiə/	'Belgium'
(3.10)	Cánada	/'kɑnada/	'Canada'

(Trommelen & Zonneveld, 1989)

Moreover, being a stress-timed language similar to English, Dutch also exhibits segmental effects of stress, particularly vowel reduction (Koopmans-Van Beinum, 1980). Hence, vowels in unstressed syllables can be reduced to schwa, but this effect is not as strong as it is in English (Kager, 1989; Trommelen & Zonneveld, 1989). The schwa itself plays a role in stress assignment in Dutch as well, as the syllable preceding a final syllable with a schwa always receives stress, unless this schwa follows another vowel immediately (see examples 3.8 versus 3.9: Kager & Zonneveld, 1986). Additionally, there are other exceptions which cannot be accounted for by QS. These exception words are reported to have lexical stress (see example 3.10).

Once the primary stress is set, secondary stress is attributed by assigning trochees and avoiding stress clash; usually, the secondary stress ends up on the initial syllable with which a so-called 'hammock' is created (van Zonneveld, 1985). Representation-wise, in the weak part of the trochaic foot (strong-weak), the vowel can optionally be reduced to schwa, but this tends to happen more often for non-high vowels in open syllables than for high vowels in closed syllables. Additionally, vowel reduction never takes place in word-final syllables. A strong-weak (SW) pattern also emerges in Dutch compounds, as the leftmost element receives the primary stress, thereby adding more weight to the fact that, statistically speaking, Dutch has initial rather than prefinal stress. At the phrase level, however, Dutch shows a different prosodic pattern which will be elaborated on in the next section.

Phrase stress

In Dutch, the rightmost element of the phrase receives phrasal prominence. According to Nespor & Vogel (1986) and Nespor, Guasti & Christophe (1996) when the relative order of head and complement constituents within the phonological phrase is head-complement, this causes the main prominence to be on the right edge of the phrase. However, in terms of word order, Dutch is a subject-object-verb (SOV) language, but also has a verb second (V2) rule for main clauses. This makes Dutch a hybrid language concerning phrasal stress, as both orders of head and complement occur in the language, similar to what has been attested for German by Nespor, Shukla, van de Vijver, Avesani, Schraudolf & Donati (2008), as illustrated in examples 3.11 and 3.12. Both left edge and right edge stress patterns thus exist at the phrase level in Dutch (complement underlined, prominence in capitals). Furthermore, the prominence in Dutch intonational phrases does not erase word stress patterns as the position of phrase level stresses determines the relative prominence of the word level stresses.

(3.11) Ik zie dat hij <u>APPELTAART</u> eet (SOV)
“I see that he apple pie eats”
‘I see that he eats apple pie’

(3.12) Hij eet <u>APPELTAART</u> (V2)
“He eats apple pie”

Phonetic realization of stress

In Dutch, a combination of pitch and intensity, spectral balance or tilt, and duration are important cues to stress (Sluijter & van Heuven, 1996a; 1996b), but intensity by itself and vowel quality, are not. Spectral balance or tilt means that intensity differences between stressed and unstressed vowels are mainly present in higher frequency bands. When considering duration, stressed syllables are longer than unstressed syllables. Both of these stress correlates are important for the perception of stress by adult Dutch listeners (Sluijter, van Heuven & Pacilly, 1997) and these may also be important for the perception of stress by Dutch-learning infants. When developing stimuli for an experimental study on the perception of stress by Dutch-learning infants, these acoustic correlates should thus be taken into account.

3.3 Turkish

Turkish (Türkçe: Underhill, 1976; Göksel & Kerslake, 2005) is part of the Turkic language family, more specifically, of the southwestern or Oghuz branch. Turkish is spoken by about 65 million speakers, of which 60 million live in Turkey. The language has several variants, but the variant spoken in Istanbul is considered to be standard Modern Turkish. This variant is also the language of which the properties are described below, because in the current study all Turkish participants were tested in Istanbul (although not all parents of the infant participants were originally from Istanbul as there are many 'immigrants' from other Turkish regions who live in Istanbul).

3.3.1 Phoneme inventory and vowel harmony

Turkish has 24 consonants: /p, b, t, d, k, g, c, ɟ, tʃ, dʒ, f, v, s, z, ʃ, ʒ, h, ɣ, m, n, ɫ, l, r, j/ and eight vowels: /a, e, i, o, u, ɯ, œ, y/. The language has vowel harmony (Kabak, 2011), a phenomenon in which the vowels within the domain of the phonological word are constrained by certain features. Vowel harmony in Turkish is determined by the word root, and spreads from left to right. It is constrained by the feature of backness in the case of the non-high vowels /a, e/, but it is additionally constrained by roundedness in the case of the high vowels /i, ɯ, œ, y/. The non-high vowels /o, œ/ are not part of the vowel harmony process, as they only occur in the initial syllable of a word, i.e. in word roots. If a word root is disharmonic, which occurs, the suffixes harmonize with the last syllable of the root (see example 3.13, for the root 'insan' meaning 'human'). There are also suffixes which are disharmonic themselves, in which case the last syllable of the suffix leads the way for further suffixation (see example 3.14 for the suffix –*iyor* for progressive tense).

(3.13)	insanlar	/insan'lar/	'humans'
		insan-lar	
		human-PL	

(3.14)	geliyorsunuz	/ge'lijorsunuz/ gel-iyor-sunuz come-PROG-2PL	'you are coming'

3.3.2 Morphology and syllable structure

Turkish is an agglutinative language, which means its morphology consists mostly of suffixes and clitics which can be added to the word root. Turkish has monosyllabic words and, in fact, most word roots are monosyllabic. The possible syllable templates in Turkish are VC, CV, CVC, VCC, CVCC and V (see example 3.15 for CVCC and example 3.16 for the only V-stem in Turkish), in order of frequency (Göksel & Kerslake, 2005). The template of CCV(C)(C) occurs only in borrowings and additionally only in the word-initial position (see example 3.17). In these borrowings, we also see loanword adaptation by speakers inserting vowels between the consonants in the word-initial clusters.

(3.15)	jant	/ʒant/	'rim'
(3.16)	o	/o/	*third person singular pronoun*
(3.17)	tren	/t(i)'ren/	'train'

3.3.3 Rhythm class

Turkish falls within the broad category of syllable-timed languages as categorized by Ramus et al. (1999) and reported on in Nespor, Shukla & Mehler (2011). According to Dauer (1983), Turkish displays many of the features of other syllable-timed languages, such as a relatively simple syllable structure and no vowel reduction. However, Turkish does not comprise all features, as fixed stress is also characteristic of many other syllable-timed languages and the Turkish stress system has exceptional non-final stress words. Thus, we cannot speak of fixed stress in Turkish, but rather of default stress (more details are given in Section 3.3.4 below).

3.3.4 Stress

Word stress

In Turkish, morphological processes play a large role in assigning stress and, for this reason, it is important to control for these processes in a study on the relationship between stress and word segmentation. As mentioned before, Turkish stress is by default word-final, but there are exceptions to this pattern concerning both word roots and suffixes. Hence, we can speak of regular and irregular stress in Turkish. On the one hand, Turkish has a simple stress assignment rule which places stress on the final syllable of a word irrespective of the morphological structure of the word and the weight of the syllables (see examples in 3.18).

(3.18)	kitáp	/ki'tap/	'book'
	kitaplár	/kitap'lar/	'books'
	kitaplarím	/kitapla'rɯm/	'my books'
	kitaplarımdá	/kitaplarɯm'da/	'in my books'
	kitaplarımdakí	/kitaplarɯmda'ki/	'the one in my books'
	kitaplarımdakilér	/kitaplarɯmdaki'ler/	'the ones in my books'
	kitaplarımdakileré	/kitaplarɯmdakile're/	'to the ones in my books'

(Göksel & Kerslake, 2005)

On the other hand, Turkish also has words with non-final stress, which can be classified into two types: exceptional roots (place names, personal names and borrowings) and exceptional suffixes, which prevent placement of stress to their right. Sezer (1983) proposed a quantity sensitivity rule to account for the stress patterns of exceptional roots, named the Sezer Stress Rule. Yet, many exceptions to this rule were still found (a.o. Inkelas, 1999, Kabak & Vogel, 2001 and Inkelas & Orgun, 2003). In fact, non-Sezer roots are more frequent among irregular roots, and all irregular roots should thus be avoided in developing stimuli for a study on default final stress in Turkish. The other type of irregularity is caused by exceptional suffix stress, for example the suffix *–iyor* mentioned earlier (see example 3.19). Kabak and Vogel (2001) propose that there is a set of suffixes and clitics which yield non-final stress within a word. These are, for example, *-iyor* (progressive aspect), which is locally accented and is thus pre-specified to carry stress on the initial vowel of the suffix (see example 3.19). Furthermore, clitics such as *–ma* (negation) create a pre-stressing effect (Example 3.20; for details see Kabak & Vogel, 2001; Kabak & Revithiadou, 2009a). These so-called 'phonological word adjoiners' all exhibit the same stress behavior: they are pre-stressing. This type of suffix should therefore be avoided when creating pseudo-word stimuli for experiments on rhythmic preferences and stress-based word segmentation in Turkish.

(3.19) gelíyor /ge'lijor/ 'coming'
(3.20) gélmesin /'gelmesin/ 'do not come'

Phrase stress
In Turkish phrases, as well as in compounds, prominence falls on the left edge of the domain, illustrated in examples 3.21 and 3.22. According to Nespor & Vogel (1986), Nespor et al. (1996) and Nespor et al. (2008), this prominence pattern at the phrase level is related to the relative order of head and complement constituents within the phonological phrase in Turkish. Turkish is a complement-head language, which causes the phrasal prominence to be on the left edge of the phrase. Furthermore, phonological phrases in Turkish are marked by boundary tones which mark the end of the phonological phrase domain by a high tone (Güneş, 2015). If words in the phrase carry exceptional non-final stress, the phrase will consequently consist of two high tones; however, in phrases with default final-stress words, the word stress and the boundary tone overlap. The overlap has been regarded as a pitch accent by Levi (2005) and Ipek & Jun (2013; 2014) and as a boundary tone by Kamali (2011).

(3.21) <u>elma</u> pasta (compound)
"apple pie"

(3.22) O <u>elmalı pasta</u> yiyor (SOV)
"he apple pie eats"
'he eats apple pie'

Phonetic realization of stress
In Turkish, stress is mostly realized by pitch, with stressed syllables having a higher pitch than unstressed syllables, and to some extent by intensity, with stressed syllables being somewhat louder than unstressed syllables (Levi, 2005), whereas duration does not play a role. These acoustic correlates of stress also match with Turkish being classified as a syllable-timed language (Schiering, 2007). The classification of Turkish stress within phonetics-based typology has been the topic of debate amongst phonologists for a long time (for the most recent 'non-stress' analysis, see Kamali, 2011) and because of its strong reliance on pitch, Turkish has also more recently been labeled a pitch-accent language, similar to Japanese (Levi, 2005). Regardless of the classification of Turkish as a stress or pitch-accent language, it is important to

control for the most important acoustic correlates of stress in Turkish in an experimental study on the perception of stress by Turkish-learning infants.

Chapter 4
Rhythmic preferences in monolingual Dutch- and Turkish-learning infants

4.1 Introduction

The youngest age at which a language-specific response related to the stress pattern of the native language has been found is 4 to 5 months (Friederici, Friedrich & Christophe, 2007). Friederici and colleagues presented 4-to-5-month-old monolingual German- and French-learning infants with two versions of the disyllabic non-word /baba/ recorded by a female native speaker of German. The stimuli were either German-like and stressed on the first (/ba:ba/) or French-like and stressed on the second syllable (/baba:/), but stress was cued by duration only. Friederici and colleagues used an oddball paradigm while measuring the infants' brain activity with electroencephalography (EEG). The analyses of the Event-related Potentials (ERPs) revealed that a mismatch negativity (MMN) was only present when the deviant stimulus carried the non-native stress pattern (for German WS and for French SW), but not the reverse.

In other words, the German-learning infants only responded to a WS stimulus in a context of SW stimuli, but not to an SW stimulus in the context of WS stimuli, and the French-learning infants showed the exact opposite pattern. It is debatable whether a durational manipulation alone can be considered a stress cue at the word level, as lengthening is mostly regarded as a universal cue to mark the end of a phrase. Furthermore, duration alone appeared not to be a reliable stress cue to word segmentation for speakers of a language similar to German, namely for Dutch-speaking adults (Lentz, Kijak & Kager, in preparation). Nevertheless, the authors viewed this result as evidence for the presence of a language-specific representation of the native language's stress pattern in 4-to-5-month-old infants. The asymmetrical discrimination effect can only be interpreted as language-specific, however this interpretation does not extend to also being evidence of a preference for the stress pattern of the native language. This is because discrimination does not automatically imply a preference, whereas reversely, a preference necessarily implies discrimination.

The step from a discrimination study using an electrophysiological method by Friederici et al. (2007) to a preference study using a behavioral method was taken by Höhle, Bijeljac-Babic, Herold, Weissenborn & Nazzi (2009). By means of the Head-turn Preference Procedure (HPP) they tested monolingual 4- and 6-month-old German- and 6-month-old French-learning infants' on their preferences for listening to an SW (native to German) version of the disyllabic non-word /'gaba/ or a WS (native to French) version of the same non-word /ga'ba/. All stimuli were recorded by a female native speaker of German. For these stimuli the cues to stress were manipulated in different ways across the stress conditions; the first syllable was always shorter and higher in pitch than the second syllable, regardless of whether the stimulus had an SW or a WS pattern. The only difference between the stress conditions was that in the SW stimuli, the difference in pitch was greater and, in the WS stimuli, the durational difference was more pronounced, thereby following the philosophy of the ITL.

Höhle and colleagues (2009) found an SW preference in German-learning infants only at 6 months of age, but not at 4 months of age, contrary to the results of Friederici et al. (2007). No preference was found in French-learning infants at 6 months of age, again in contrast with the language-specific discrimination results found by Friederici and colleagues. However, Höhle et al. (2009) do show general discrimination of the different stress patterns by French-learning infants at 6 months of age using the HPP, indicating that the lack of a preference is not due to a failure to discriminate between the stress patterns. As the very early language-specific results for both German- and French-learning infants found by Friederici et al. (2007) have not been replicated (behaviorally) we need to be cautious in concluding that German-learning infants have an SW preference at 4-5 months of age and that French-learning infants possess language-specific knowledge of the dominant rhythmic pattern of their native language at this early age.

Upon review, both of the above mentioned studies investigated the processing of stress in the absence of segmental variation, as these studies present infants with multiple tokens of a single non-word. However, there are differences in the age at which a rhythmic preference is found for these and other studies using a different type of stimuli. These differences may be due to the amount of variation in the stimuli used, an issue I will discuss more elaborately. Jusczyk, Cutler & Redanz (1993a) explored the processing of stress information in the presence of segmental variation by 6- and 9-month-old monolingual English-learning infants, in a study using the HPP. They found that infants listened longer to lists of low-frequent English SW words than to lists of low-frequent English WS words at 9 months of age, but not yet at 6 months. They inferred from this that a language-specific rhythmic preference develops between 6 and 9 months of age. Moreover, at 9 months the effect was maintained when presenting infants with the low-pass filtered version of the stimuli. This demonstrates that the SW preference at 9 months can also be elicited by presenting only the prosodic features of the stimuli.

Unfortunately, this study did not test the 6-month-old infants with a low-pass filtered version of the stimuli. In retrospect, this would have enabled a comparison with the studies discussed above using a single non-word, i.e. stimuli without segmental variation, as these studies did find native preferences to emerge as early as 6 months of age. The differences between the results from different age groups and experiments using different types of stimuli imply that processing stress in the absence of segmental variation is less difficult. In summary, the findings from these cross-linguistic rhythmic discrimination and preference studies are only partly congruent. Friederici et al. (2007) found language-specific discrimination of stress patterns for both German- and French-learning infants at only 4-5 months of age. Yet, Höhle et al. (2009) failed to find a language-specific rhythmic preference for German-learning infants at 4 months of age, but did find it at 6 months of age. For French-learning infants at 6 months of age, they did not find any preference, although they did find evidence of discrimination between the SW and WS versions of a single non-word. This discrepancy between discrimination and preference studies means that for both German- and French-learning infants the language-specific discrimination found by Friederici et al. (2007) at 4-to-5 months of age cannot be supported by evidence from the behavioral preference experiment by Höhle et al. (2009) with 4- and 6-month-old German- and French-learning infants, respectively. As we are dealing with experiments using different procedures, electrophysiological and behavioral, and different stimuli, using only duration as a cue to stress or pitch only for the SW items and duration only for the WS items, these factors may be the cause of the incongruency. Additionally, the study by Jusczyk et al. (1993a) confirms that processing stress information on segmentally varied stimuli may be more difficult than processing this information in the absence of this variation.

Besides these methodological issues, the differences between the languages tested may underlie the different results found for German-, English- and French-learning infants. According to Nazzi et al. (1998) learners of stress-timed languages such as English, German and Dutch start segmenting the speech stream on the basis of a disyllabic SW rhythmic unit, while learners of syllable-timed languages such as French and Turkish start segmenting the input on the basis of a syllabic unit. This is one reason why French may not be an appropriate language for testing the development of rhythmic preferences on disyllabic stimuli. Additionally, and contrary to what is occasionally assumed, French does not use stress at the word level, but only uses a phrase final 'accent' (Grammont, 1965; Dupoux et al., 1997; Jun & Fougeron, 2000).

Höhle et al.'s study suggests that the lack of a rhythmic preference in 6-month-old French-learning infants may be explained in one of two ways: (i) either infants learning a syllable-timed language, such as French, do not process speech in rhythmic units, but in syllabic units, or (ii) infants learning a language which does not have stress at the word level, such as French, do not use stress as a cue to word well-formedness and, therefore, they do not develop a rhythmic preference. I would like to add a third hypothesis here, namely (iii) the existence of a universal trochaic bias in acquisition (Allen & Hawkins, 1978; 1980), which is only 'activated' by the presence of word level stress in the language being learned. Further investigation of this hypothesis is warranted as perceptual preferences in acquisition have been found for Germanic trochaic languages, which use stress at the word level, only. These 'language-specific' SW preferences can therefore not be disentangled from a universal trochaic bias, but may suggest that a trochaic bias is only developed in infants learning a language which has stress at the word level, rather than at the phrasal level.

To disentangle the three hypotheses mentioned above, we have to test a language which is syllable-timed, but has a WS stress pattern at the word level. Turkish is one such language. Like French, it is syllable-timed, but unlike French, it has final stress at the word level. Turkish has default stress on the last syllable of the word, which often is a suffix due to the agglutinative character of the Turkish language. Turkish, however, has exceptions to this final stress pattern, mostly in proper names and under certain morphological conditions, such as pre-stressing suffixes, which also distinguishes the Turkish system from French (Göksel & Kerslake, 2005). In light of these properties, Höhle et al. (2009) are justified in concluding that in order to find out which of the first two mentioned hypotheses holds, infants learning Turkish should be tested. If indeed Turkish has word-final stress I expect this language to also be an adequate test case for the third hypothesis, the UTB hypothesis.

A previous attempt at testing Turkish-learning infants on their rhythmic preferences was made by van Kampen, Parmaksiz, van de Vijver & Höhle (2008). They used the HPP to test Turkish-learning infants born in the Turkish community in Berlin with, on average, 2.4 hours of German input per day (with a range of 0-10 hours), and monolingual German-learning infants at 6 months of age. The stimuli consisted of different disyllabic CVCVC non-words with either stress on the first or the second syllable. The stimuli were recorded by a native speaker of German for the German-learning infants and a native speaker of Turkish for the Turkish-learning infants. The results showed no effect of stress pattern (SW vs WS) for either group. The results for the German group were surprising in that they were in contrast to previous results from German-learning 6-month-olds. However, before further assessing this result, it should be noted that word stress was not a within-subjects factor, but a between-subjects factor in this study. In other words, the differences in listening times measured were always across different groups of infants and are therefore not directly comparable. The lack of an effect for the factor stress pattern may either be due to the fact that processing stress information over stimuli with segmental variation is too difficult for infants of this age (see Jusczyk et al., 1993a), or by its status as a between-subjects factor.

In order to evaluate rhythmic preferences in Turkish-learning infants properly, infants should be presented with stimuli without segmental variation and where stress pattern is a within-subjects factor. In such an experiment, note should be taken of another issue; in the previously mentioned cross-linguistic experiments (Friederici et al., 2007 and Höhle et al., 2009) the stimuli were not controlled for native listening in both of the language groups under consideration. In these cross-linguistic experiments comparing German- and French-learning infants all stimuli used were those of a native speaker of German, creating a possible advantage for the German-learning infants in processing them. The stimuli were neither controlled for the segmental properties of the languages involved, especially not in interaction with stress, nor were the stimuli properly controlled for the relevant language-specific phonetic cues to stress. Friederici et al. (2007) only used duration as a cue to stress and the stimuli used by Höhle et al. (2009) did not contain unambiguous cues to initial and final stress, as the stressed syllable was not always longer in duration and higher in pitch than the unstressed syllable, considering the first syllable was always shorter and higher in pitch than the second syllable.

To disentangle the aforementioned hypotheses: (a) the rhythm class account, (b) the word level versus phrase level account and (c) the universal trochaic bias account, a cross-linguistic preference experiment was designed for testing two languages from different rhythmic groups, namely Dutch and Turkish. The focus of this study was on Turkish with the specific aim of finding out whether Turkish-learning infants between 4 and 8 months of age, unlike French-learning infants, show either i) a WS preference based on the dominant word-final stress pattern of their native language, (ii) do not show any preference similar to French-learning infants, or (iii) alternatively show a universal trochaic bias. The experiment for Dutch was a baseline experiment with the aim of exploring whether Dutch-learning infants between 4 and 8 months of age, like German-learning infants, show an SW preference at 6 months of age. Dutch is a stress-timed language with a statistically dominant SW stress pattern in disyllabic words and it is also described as having a trochaic disyllabic foot (Kager, 1989; Trommelen & Zonneveld, 1989), similar to German and English.

Because of the large age range between 4 and 8 months of age, I will also analyze the results by investigating the effect of age group on the development of familiarity and novelty preferences during the experiment (Hunter & Ames, 1988). This methodological aspect has often been overlooked in earlier studies testing different age groups with the same experiment. Longer Looking Times (LTs) during one of the conditions is generally interpreted as a familiarity preference without taking into account factors such as age and task complexity. All infants were tested in the same paradigm, namely a Central Fixation Auditory Preference Procedure based on Cooper & Aslin (1990), with the use of eye tracking. Infants were presented with stimuli created to avoid the pitfalls discussed above. The stimuli lacked segmental variation and possessed all of the relevant phonetic cues to stress used in both languages: pitch, duration and intensity. The stimuli were also controlled for the interaction of stress placement with segmental properties for both languages and were recorded by a speaker of a third language in order to avoid a processing advantage for one of the language groups being tested in the current cross-linguistic study.

4.2 Experiment 1: method and results

4.2.1 Participants

The participants in this experiment were 102 Dutch-learning infants: 31 infants around 4 months of age (M = 139 days), 41 infants around 6 months of age (M = 203 days), and 30 infants around 8 months of age (M = 262 days) tested in a cross-sectional design. All participants were healthy full-term infants, without any known hearing or visual impairments, and were being raised in a monolingual Dutch environment. In the 4-, 6- and 8-month-old groups another 13, 10 and 13 infants were tested, respectively. Data from these infants could not be included due to crying (15), fussiness (11) or technical issues (10). The participant information is summarized in Table 4.1.

Table 4.1: *Included and total number of Dutch-learning infants tested per age group with mean age and age range in days.*

Age group	*4-month-olds*	*6-month-olds*	*8-month-olds*
Included (total)	31 (44)	41 (51)	30 (43)
Mean age (range)	139 (125-160)	203 (185-220)	262 (238-283)

4.2.2 Material

The infants were presented with an auditory stimulus, the disyllabic CVCCVC sequence /noldaf/, forming a pseudo-word in both Dutch and Turkish. This pseudo-word is phonotactically well-formed in both languages under investigation and stressed on either the first or the last syllable, forming an SW or WS unit, respectively. The pseudo-word was controlled for segmental and phonotactic naturalness by using phonemes with similar frequencies and transitional probabilities (TPs) in Dutch and Turkish. More specifically, the phoneme inventories, the TPs of the adjacent consonants, the triphone frequencies of both CVC syllables, the cohort sizes of the first syllable, and the syllable frequencies of each syllable in combined initial versus final position and stressed versus unstressed position were taken into account. In addition, the stimulus was also controlled for vowel harmony, which means the pseudo-word is harmonious for vowel backness. This is relevant for Turkish, as Turkish words must adhere to the vowel harmony rules of the language (Göksel & Kerslake, 2005). All of these frequencies were calculated by analyzing the phonemically transcribed Dutch CELEX version 3.1 corpus (Baayen, Piepenbrock & Gulikers, 1995) and the Turkish TELL version 2.0 corpus (Inkelas, Küntay, Orhan Orgun & Sprouse, 2000) using PhonotacTools (Adriaans, 2006).

Twelve tokens for each stress pattern were recorded by a female native speaker of Spanish, a foreign language to both language groups. The decision to involve a third language speaker was made in order to avoid an advantage for either the Dutch or Turkish language group. Spanish was chosen, because this language has lexical stress, which means the stress pattern of words is lexically specified, so that producing stress in different positions of the word was a relatively easy task for the speaker. The speaker was instructed to produce the pseudo-word in a Spanish carrier sentence and in an infant-directed manner. As the speaker is a mother herself, doing this was also relatively easy for her. This methodological decision is different from those made in earlier studies reviewed here and I will return to this decision to use a third language speaker in the discussion section.

The pseudo-words were cut from their carrier sentence and subjected to an acoustical analysis using PRAAT (Boersma & Weenink, 2011). In this analysis, the usual components of duration, pitch and intensity on the rhyme part of each syllable were measured from the spectrogram and then submitted to a statistical analysis, in order to ensure the right stress cues were present in the recordings. According to Levi (2005), stress in Turkish is mostly realized by pitch. In Dutch, a combination of pitch and intensity (spectral tilt) and duration are important cues to stress (Sluijter & van Heuven, 1996a; 1996b; Sluijter et al., 1997). As presented in Table 4.2, stressed syllables (meaning the rhyme portion of the syllable as this is the domain of syllable duration) had an average duration of 289 milliseconds (ms), an average mean pitch of 348.9 Hz, and an average mean intensity of 70.4 dB, whereas unstressed syllables had an average duration of 225 ms, an average mean pitch of 198.3 Hz, and an average mean intensity of 63.3 dB. A paired samples t-test (2-tailed) revealed that the stressed syllables (always only the rhyme portion) had a significantly longer duration (difference: 64 ms; $t(23) = 5.08$, $p<.001$), a significantly higher mean pitch (difference: 150.6 Hz; $t(23) = 20.03$, $p<.001$) and a significantly higher mean intensity (difference: 7.1 dB; $t(23) = 15.71$, $p<.001$) than (the rhyme portion of) the unstressed syllables.

In addition to the acoustical and statistical analyses the stimuli were judged by adult native speakers from both language groups. The participants of this assessment task were selected as monolinguals and were asked to judge each stimulus on the position of the stressed syllable, that is, to answer the following question: was the stress placed on the first or on the last syllable? 10 monolingual adult native speakers of Dutch judged the stress placement 100% correctly and 3 monolingual adult native speakers of Turkish judged the stress placement 79% correctly (not at ceiling, which is not unexpected given the findings of Domahs, Genc, Knaus, Wiese & Kabak, 2012a). 10 out of the 12 recorded tokens per stress pattern were selected, based on the correct stress placement judgements by all Turkish native speakers. These tokens formed 10 trials per stress condition each containing 10 tokens with the same stress pattern in a pseudo-randomized order within the 10 trials. The other 2 tokens per stress pattern were used to create a short trial of 6 seconds which was used to familiarize the infants with the auditory and visual stimuli. Within each test trial there was an inter stimulus interval (ISI) of 750 ms between each token, to prevent an iterative pattern being perceived on more than two syllables. Due to the fact that the first syllable of the pseudo-word inherently had a longer duration than the last syllable regardless of whether it was stressed or unstressed, the SW trials turned out to take slightly longer than the WS trials. The SW trials had a mean duration of 15424 ms, ranging between 15280 and 15440 ms, whereas the WS trials had a mean duration of 14744 ms, ranging between 14720 and 14880 ms. As the experiment used fixed trials and a different number of tokens per trial between the conditions would have been undesirable, the difference in duration was corrected for in the data analysis and this correction will be described in the results section.

Table 4.2: *The average values over 24 tokens for duration, mean pitch and mean intensity of the rhyme part of the stressed and unstressed syllables with their standard deviation (SD).*

Syllable type	*Duration (ms)*	*Pitch (mHz)*	*Intensity (mdB)*
Stressed (SD)	289 (38.6)	348.9 (36.5)	70.4 (2.0)
Unstressed (SD)	225 (41.8)	198.3 (9.2)	63.3 (1.8)

4.2.3 Procedure

Rather than using the traditional HPP, a Central Fixation Auditory Preference Procedure (Cooper & Aslin, 1990) by means of eye tracking was used to measure rhythmic preferences. Firstly, in the HPP, looking times are measured online by a human observer who scores the head or eye movements of the infants. This method has poor inter-observer reliability caused by unwanted variation (noise) within experiments when different observers score different infants within the same experiment, but also between experiments aiming to study the same phenomena in different labs which may use different scoring criteria. Eye tracking, on the other hand, provides us with a more objective measure of eye movements, as these are scored by a computer, using the same criteria both within and between experiments, when provided with the same settings (Wass, Smith & Johnson, 2013). Secondly, it is known that infants under the age of 5 months may still have some difficulty disengaging from a stimulus in the HPP (Frick, Colombo & Saxon, 1999). This difficulty may be less prominent in a paradigm in which they do not have to move their heads to indicate disinterest, rather disinterest is indicated through a subtle movement of the eyes. A final consideration for this cross-linguistic study was that the methodology used was available in both language environments under investigation and the HPP set-up was not available in Turkey at the time of testing.

Hence, in this part of the study a Tobii 1750 eye tracker with so-called Infant Add-on illumination – an extra piece of hardware used for tracking infants (Tobii Technology AB, 2006) – was connected to a Paradigit MSI 945P Platinum computer, and was used at the baby lab of the Utrecht Institute of Linguistics OTs, part of Utrecht University in the Netherlands. The eye tracker was placed inside a sound-proof booth in which the caregiver was seated on a chair with the infant on his or her lap in a special car seat aligned with the (moveable) monitor of the eye tracker. During the experiment the caregiver listened to music over headphones and was instructed not to interact with the infant. The infant faced the 17” screen of the eye tracker at a distance of approximately 25.6 inches (or 65 cm). Directly below the screen a Fostex PM0.4 active loudspeaker was placed, through which the auditory stimuli were played at a comfortable loudness.

The experiment was run in and recorded with ClearView analysis software version 2.7.1. The auditory and accompanying visual stimuli were played as a continuous AVI movie. The experiment started with an attention grabber (Johnson in Tobii Technology AB, 2006) showing a moving duckling accompanied by a ringing sound, which was followed by a short familiarization trial of 6 seconds in which the infants heard four alternating tokens of the pseudo-word (two tokens per stress pattern which were different from the tokens used in the test trials) whilst looking at the visual stimulus presented during all trials in the experiment. The visual stimulus was a static picture of a smiling female face (Figure 4.1). After the familiarization trial (which was not included in the data analysis) and after each test trial the attention grabber appeared again to ensure the infant was looking at the screen at the beginning of the next test trial.

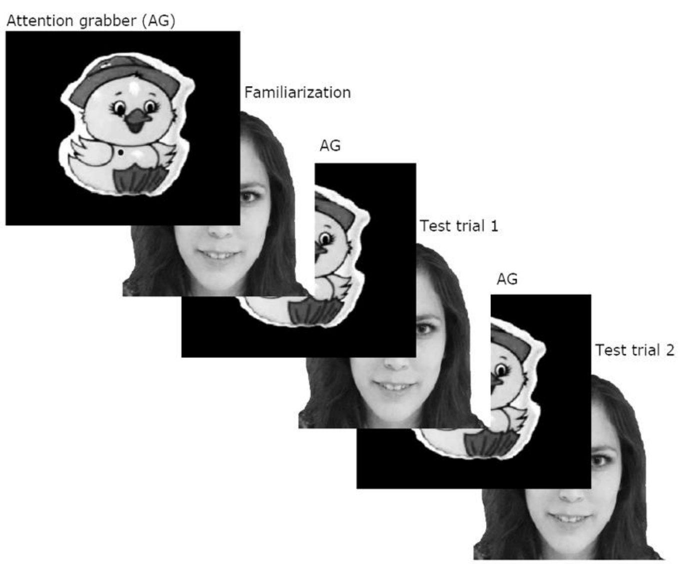


Figure 4.1: *The outline of the preferential listening experiment using eye tracking.*

The test phase consisted of 20 alternating trials; 10 per stress condition, either starting with an SW or a WS trial (counterbalanced), accumulating to an experiment of 8 minutes and 20 seconds in total. However, before starting the actual experiment, the eye tracker had to be calibrated for each infant in order to be able to correctly track their eyes during the experiment. A five-point-calibration was performed with ClearView 2.7.1 software and consisted of presenting the same attention grabber as in the experiment at five points on the screen, in all four corners and the middle of the screen. The infant is supposed to follow the attention grabber and focus on all five points, so that the software can make a representation of where the eyes are situated relative to the screen for each individual infant. If the calibration process, after several re-tries, was still unsuccessful the infant could not proceed to the actual experiment which was, however, rarely the case. When calibration was successful, the infant proceeded to the experiment and the recording ended when the infant stopped attending completely, due to fussiness or crying, or when the experiment finished.

4.2.4 Data analysis

In order to extract the information needed to calculate the Total Looking Time (TLT) per participant per trial the temporal and spatial criteria for 'looking' and 'looking away' were determined. The Tobii 1750 extracts a gaze sample approximately every 20 ms (at a sampling rate of 50 Hz) providing the X- and Y-coordinates of the location of the eye gaze on the screen. In defining 'looking', I decided to not only include fixations (classified as the samples where gaze remained on the same location of the screen for a minimum of 200 ms), but also all gaze samples longer than 20 ms captured by the eye tracker within the dimensions of the screen. This means that all gaze samples larger than 20 ms located outside the dimensions of the screen were coded as 'looking away'. I chose the screen as the spatial domain because it is comparable to what has been done in similar methods such as the Central Fixation Auditory Preference Procedure (Cooper & Aslin, 1990) during which a human observer decides whether the infant is looking at the screen or not.

In determining the inclusion criteria for trials, it was decided that only completed trials were included in the analysis. In other words, if the experiment was stopped during a trial, this trial would be excluded from the analysis. The participant inclusion criteria stipulated that only infants who completed at least half of the experiment, that is, five or more trials per stress condition, were to be included in the analysis. This criterion serves to ensure the reliability of the sample, i.e. enough data points per infant and only of infants who in general paid attention to the experiment. This criterion is similar to one of the inclusion criterion in Altvater-Mackensen, Van der Feest & Fikkert (2014). Finally, the longer duration of the SW trials was corrected for by reducing all trials that had longer TLTs than 14720 ms (the duration of the shortest WS trial) to this particular duration similar to Höhle et al. (2009).

4.2.5 Results

The purpose of Experiment 1 was to see whether Dutch-learning infants between 4 and 8 months have a preference for the SW stress pattern of their native language, and if so, at which age this preference develops. For the Dutch-learning infant cohort, the mean TLT during the SW condition was 7483 ms (SD = 3759 ms), and the mean TLT during the WS condition was 7193 ms (SD = 3627 ms). These mean TLTs per condition are presented in Figure 4.2. The mean TLTs per age group are shown in Table 4.3 and Figure 4.3.

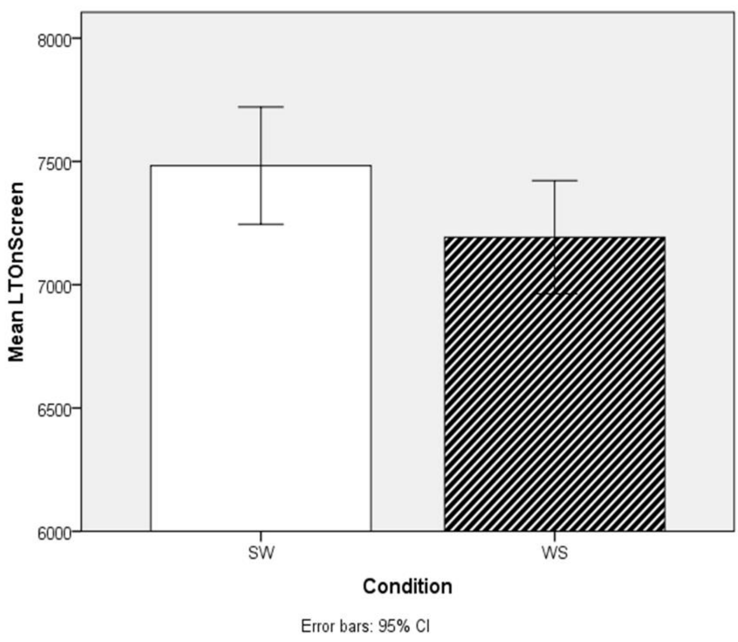


Figure 4.2: *The mean TLT in ms for all Dutch-learning infants per condition (SW vs WS).*

Table 4.3: *The mean TLT in ms per Dutch age group (4-, 6-, 8-month-olds) per condition (SW vs WS).*

Age group	*4-month-olds*	*6-month-olds*	*8-month-olds*
SW condition (SD)	7960 (3945)	7264 (3682)	7269 (3621)
WS condition (SD)	7816 (3844)	6764 (3424)	7118 (3578)

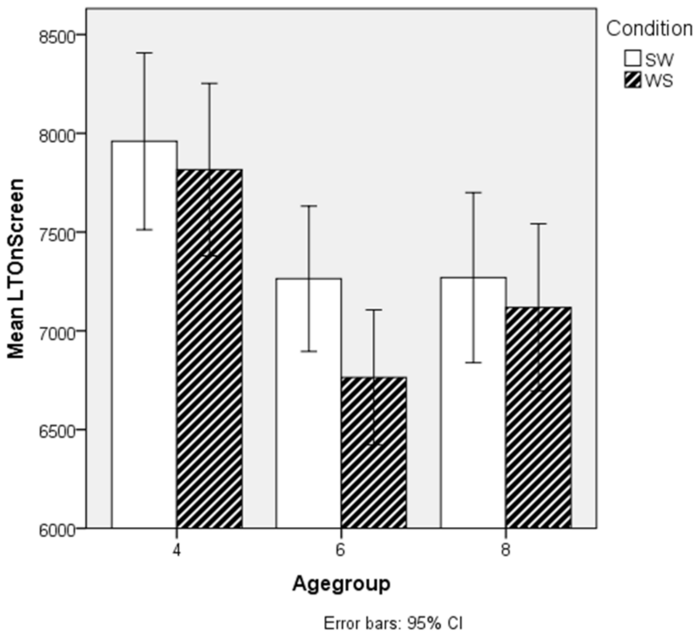


Figure 4.3: *The mean TLT in ms per Dutch age group (4-, 6-, 8-month-olds) per condition (SW vs WS).*

A statistical analysis was ran In order to determine whether there were significant effects of the variables *condition* and *age group*. A normal distribution is a prerequisite for most statistical analyses, but the data were not normally distributed. For this reason, I first ran a normal transformation on the data using Blom's formula (Blom, 1958) and then ran a statistical analysis on the normalized TLTs. Statistical analysis was done by means of Linear Mixed Model analyses, a form of General Linear Model that does not assume homogeneity of variance, sphericity or compound symmetry, and, moreover, allows for missing data (Quené & van den Bergh, 2008; Goldstein, 2011; Snijders, 2011). I ran a Linear Mixed Model analysis with condition and age group as fixed factors, participant and trial as random factors and trial order as a covariate. I found a significant main effect of the variable condition ($F(1,1815) = 7.307$, $p < .01$), with a longer mean TLT during the SW trials than during the WS trials. This means that the Dutch-learning infants, with all age groups combined, show a preference for listening to the SW stimuli over the WS stimuli. However, there is neither a significant main effect of the variable age group ($F(1,1815) = 2.123$, $p = .125$), nor a significant interaction of the variables age group and condition ($F(1,1815) = 0.777$, $p = .460$). This means the age groups behaved similarly in general and also with respect to the variable condition. Therefore, I cannot confirm a development in rhythmic preference between 4 and 8 months of age.

Nevertheless, as the set-up of the study is developmental and the design is cross-sectional, planned analyses for each age group separately are possible. When we look at the effect sizes (ES in Cohen's d) for the factor condition within each age group, the ES is largest in the 6-month-old age group (0.140), and much smaller in the 4- and 8-month-old age groups (0.036 and 0.041, respectively). These results suggest that for Dutch-learning infants, the most pronounced SW preference is at 6 months of age and less so at 4 months of age. However, the Dutch-learning infants also do not seem to show a strong preference at 8 months of age, which is unexpected as they seem to demonstrate a preference for the stress pattern of their native language by 6 months of age. As can been seen from the mean TLTs per condition in Table 4.3 and from Figures 4.2 and 4.3, the differences in mean TLTs between the conditions are rather small in general. This can possibly be explained by two factors, which I will elaborate on in the discussion in Section 4.6, after discussing the results from Experiment 2 examining the rhythmic preferences of Turkish-learning infants.

4.3 Experiment 2: method and results

4.3.1 Participants

The participants in this experiment were 92 Turkish-learning infants: 30 infants around 4 months of age (M = 140 days), 31 infants around 6 months of age (M = 195 days), and 31 infants around 8 months of age (M = 256 days) in a cross-sectional design. All participants were healthy full-term infants, without any known hearing or visual impairments, and were being raised in a monolingual Turkish environment. In the 4-, 6- and 8-month-old group another 10, 11 and 21 infants were tested, respectively. Data from these infants could not be included due to crying (15), fussiness (12), technical issues (9) or parental interference (6). The participant information is summarized in Table 4.4.

Table 4.4: *Included and total number of Turkish-learning infants tested per age group with mean age and age range in days.*

Age group	*4-month-olds*	*6-month-olds*	*8-month-olds*
Included (total)	30 (40)	31 (42)	31 (52)
Mean age (range)	140 (117-160)	195 (179-219)	256 (238-279)

4.3.2 Material

The materials used were identical to those used in Experiment 1.

4.3.3 Procedure

The procedure was largely the same as in Experiment 1, the only difference was the use of an alternative eye-tracker, namely a Tobii T120 at Koç University in Istanbul, Turkey. Both eye trackers used were suitable for the experiments, however the Tobii T120 ran at an increased sampling rate of 120 Hz, extracting a gaze sample approximately every 8 ms. The eye tracker was placed inside a sound-attenuated room at the university's Language and Communication Development Lab in the Psychology department. The calibration and the experiment were both run in and recorded with Tobii Studio analysis software version 2.3 on a Dell Precision T5500 computer connected to the eye tracker.

4.3.4 Results

The data were analyzed in the same way as the data from Experiment 1. For all age groups combined the mean TLT during the SW condition is 7602 ms (SD = 4213 ms) and the mean TLT during the WS condition is 7138 ms (SD = 3993 ms) as can been seen in Figure 4.4. The mean TLTs per age group are presented in Table 4.5 and Figure 4.5.

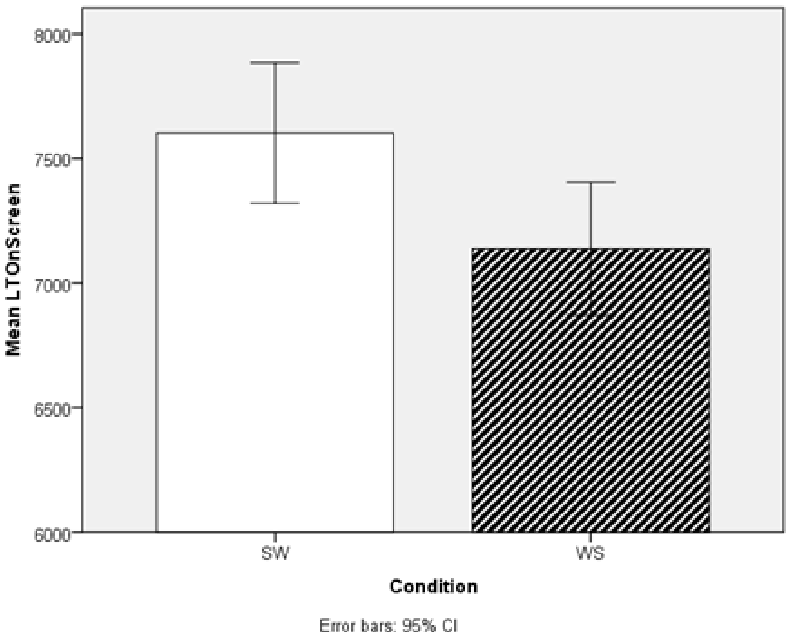


Figure 4.4: *The mean TLT in ms for all Turkish-learning infants per condition (SW vs WS).*

Table 4.5: *The mean TLT in ms per Turkish age group (4-, 6-, 8-month-olds) per condition (SW vs WS).*

Age group	*4-month-olds*	*6-month-olds*	*8-month-olds*
SW condition (SD)	7522 (4136)	7608 (4325)	7675 (4183)
WS condition (SD)	6729 (3860)	7292 (3970)	7380 (4124)

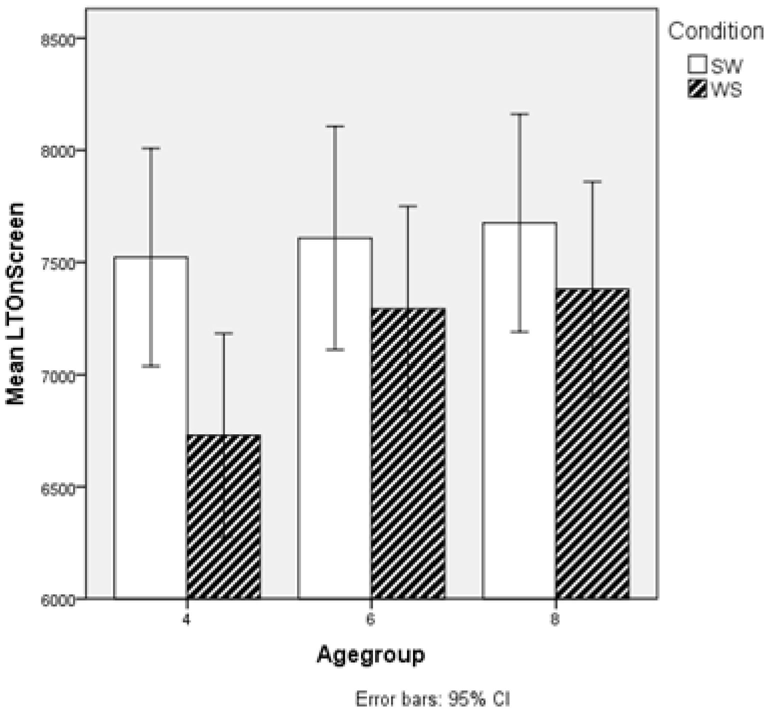


Figure 4.5: *The mean TLT in ms per Turkish age group (4-, 6-, 8-month-olds) per condition (SW vs WS).*

A significant main effect of the variable condition was found ($F(1,1619) = 11.851$, $p < .01$), with a longer mean TLT during the SW trials than during the WS trials. This means that the Turkish-learning infants, with all age groups combined, show a preference for listening to the SW stimuli over the WS stimuli. However, there is neither a significant main effect of the variable age group ($F(1,1619) =.128$, $p = .880$), nor a significant interaction of the variables age group and condition ($F(1,1619)= 1.460$, $p=.233$), which means the age groups behave similarly in general and also with respect to the variable condition, similar to the Dutch-learning infants. Therefore, I cannot confirm a development in rhythmic preference between 4 and 8 months of age for the Turkish-learning infants either. The preliminary results of this traditional analysis of the data (comparing group means across conditions) require a more in-depth analysis of the development of the infants' preferences during the experiment.

It is interesting to further examine more precisely what is occurring within each age group separately, as the difference in mean TLTs between the conditions seems to differ for each of the age groups. When we look at the ESs for the factor condition within each age group in the Turkish language group, the ES is largest in the 4-month-old age group (0.198) and much smaller in the 6- and 8-month-old age groups (0.076 and 0.071, respectively). These results suggest that the SW preference is most pronounced at 4 months of age and less so at 6 and 8 months of age. Nevertheless, as can been seen from the mean TLTs per condition in Table 4.5 and from Figures 4.4 and 4.5, the differences in mean TLTs between the conditions are again generally rather small. This can possibly be explained by two factors, namely individual differences and temporal information, which will both be discussed below.

4.4 Discussion

Firstly, by looking at the mean TLTs per condition per age group, all individual differences are collapsed. In other words, we do not know how many infants actually show a preference for the SW pattern in the Dutch-learning 6-month-old age group, and we are also unsure whether all infants in the Dutch-learning 4- and 8-month-old age groups do not show any preference, or whether there is a different pattern visible in the data. The same holds for the Turkish-learning infants: we do not know how many infants show a preference for the SW pattern in the 4-month-old age group and we cannot see whether all infants in the 6- and 8-month-old age groups do not show any preference, or whether we can see a more mixed pattern of behavior in these age groups.

When we examine the individual differences in looking times between the mean TLT during the SW and WS condition per individual infant per age group per language group, we see a pattern emerge. The distributions are represented in Figure 4.6 for Dutch and in Figure 4.7 for Turkish. In these figures, the horizontal axis is age in days and each square, circle or triangle (for 4-, 6- and 8-month-old age groups, respectively) represents the difference score in ms of an individual infant. A mean negative score always represents a familiarity preference, which means an SW preference for the Dutch-learning infants and a WS preference for the Turkish-learning infants, while a mean positive score always represents a novelty preference, meaning a WS preference for the Dutch-learning infants, but an SW preference for the Turkish-learning infants. What is apparent on the vertical axis is the range of individual preferences: from ‘no preference’ around the zero line, towards both ‘familiarity’ going negative, and ‘novelty’ preferences going positive, in all age groups and both language groups.

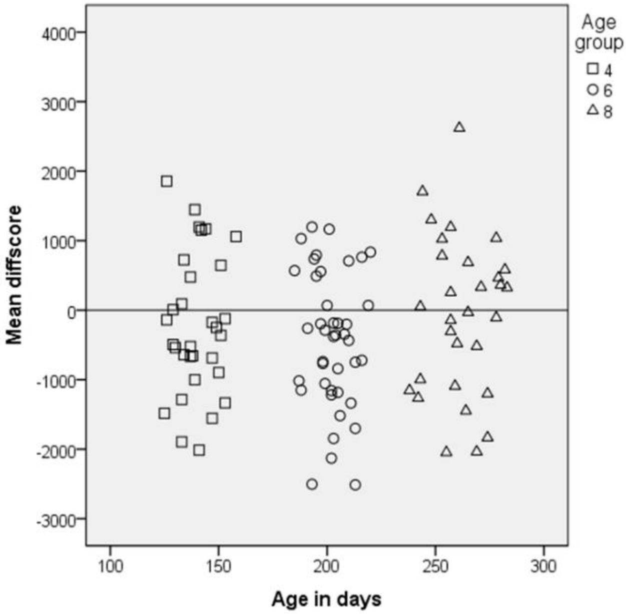


Figure 4.6: *The distribution of individual preferences by difference score in ms in the Dutch group per infant, per age group.*

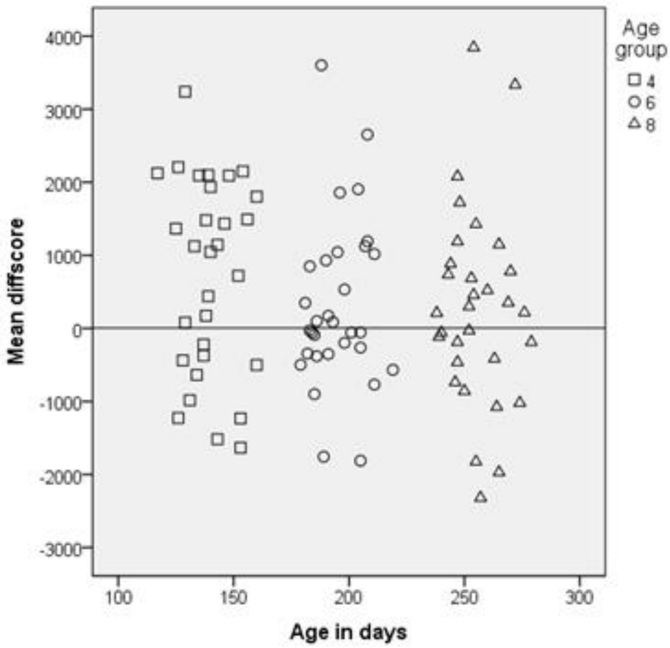


Figure 4.7: *The distribution of individual preferences by difference score in ms in the Turkish group per infant, per age group.*

In the Dutch language group there are, proportionally speaking, more infants with an SW, familiarity preference (62% mean negative scores), although there are also quite a number of infants with a WS, novelty preference (38% mean positive scores), adding up to a modest effect in the direction of an SW preference for the Dutch language group as a whole. We see a similar trend but in the opposite direction for the Turkish language group: most of the infants in this group show an SW novelty preference (58% mean positive scores), whereas there is also a relatively large group showing a preference in the opposite direction (42% mean negative scores), namely a WS familiarity preference. These individual differences between infants, cancelling out a large group effect in one direction, might be contributing to the relatively small differences found in mean TLTs between the stress conditions.

Secondly, the relatively small differences between the conditions may also be due to collapsing the mean TLTs over trials per condition through which all temporal information of trial order per condition is lost. This temporal information reveals how the direction of the infants' preferences develops over the course of the experiment. This temporal analysis is important because it can clarify what the direction of the infants' preferences actually means. Assuming the dominant stress pattern of the infants' native language is most familiar to them, we can interpret a preference of the Dutch-learning infants for the SW pattern as a familiarity preference, similar to the interpretation of the preference for the SW pattern by German- and English-learning infants in previous studies (recall the introduction section of this chapter). This means that infants who show a preference for the WS pattern are then interpreted to have a novelty preference instead of a familiarity preference. For the Turkish-learning infants the assumption is as follows: a preference for the WS pattern can be interpreted as a familiarity preference and a preference for the SW pattern as a novelty preference, based on the idea that the dominant stress pattern of their native language is most familiar to them.

Predicting under which conditions familiarity or novelty preferences will occur and how they may develop over the course of an experiment is a complex undertaking. One of the first attempts at making such predictions was made by Hunter & Ames (1988). Their model has its roots in optimal-level theories (first mentioned by Dember & Earl, 1957, but for more recent work this topic on infants see Aslin, 2007 and Kidd, Piantadosi & Aslin, 2010; 2012; 2014). Optimal-level theories state that there is an optimal level of stimulation that makes stimuli that are below or above this optimum point less interesting. According to Hunter & Ames (1988) the validity of optimal-level theories can be tested by comparing "the amount of exploration elicited by a familiarized stimulus with the amount elicited by a novel stimulus (…) in a preference test between the stimulus that is being familiarized and a novel stimulus at different points during familiarization."

Fortunately, conducting this 'preference test' is possible on the data gained in these experiments. For the Dutch-learning infants, the familiarized stimulus is the SW condition and the WS condition is the novel stimulus. For the Turkish-learning infants, the WS condition is the familiar stimulus, while the SW condition is the novel stimulus. The TLTs during these conditions can be compared by calculating the difference between the TLTs during the SW and WS trials per trial pair. This means we will get 10 difference scores (in ms) in total, one for each pair of trials, which will tell us whether the infants show no preference, a familiarity preference, or a novelty preference at that particular point in the experiment. Hunter & Ames (1988) predict that preferences should occur in a particular sequence: initially no preference, then a familiarity preference, followed by a brief period of no preference when shifting to a novelty preference. This development of preferences over time is depicted in Figure 4.8.

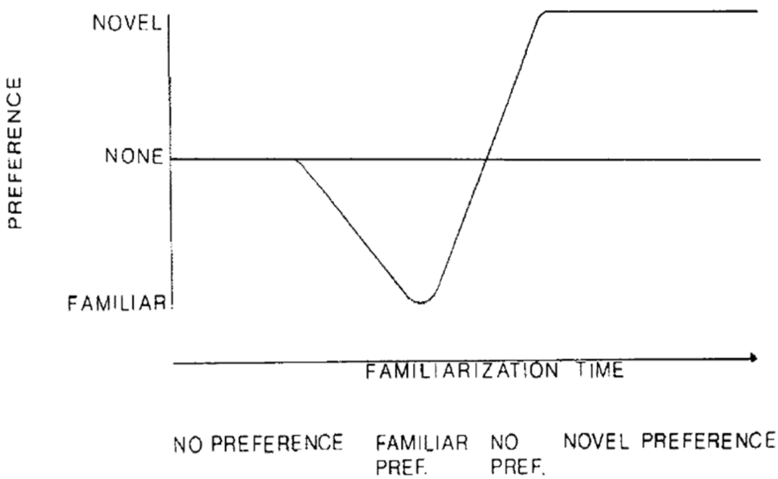


Figure 4.8: *The basic preference curve over time as described by Hunter & Ames (1988), starting with no preference, then a familiarity preference, followed by a brief period of no preference when shifting to a novelty preference.*

According to the authors, the shape of the basic preference curve cannot be altered, more specifically, the order of the preferences will always be the same. However, the amount of time it takes to complete the different preference stages can be affected by two main variables: age or cognitive maturity, and task difficulty or cognitive complexity. For the variable age, the idea is that older infants habituate faster to the familiar stimulus due to a more advanced cognitive maturation and therefore show an earlier shift to novelty than younger infants. Task difficulty, on the other hand, can slow down the rate at which the familiarity-to-novelty preference sequence is completed; when the task is cognitively more complex a later shift to novelty can be observed. In my experiment, the variable task difficulty was held constant, whereas the variable age differed between the 4-, 6- and 8-month-old age groups. Consequently, I expect to observe differences in the duration of the preference stages between the age groups.

The preference development for all three age groups and both language groups are plotted in Figure 4.9 for Dutch and in Figure 4.10 for Turkish. Again, a mean negative score always represents a familiarity preference, which means an SW preference for the Dutch-learning infants and a WS preference for the Turkish-learning infants, while a mean positive score always represents a novelty preference, meaning a WS preference for the Dutch-learning infants, but an SW preference for the Turkish-learning infants.

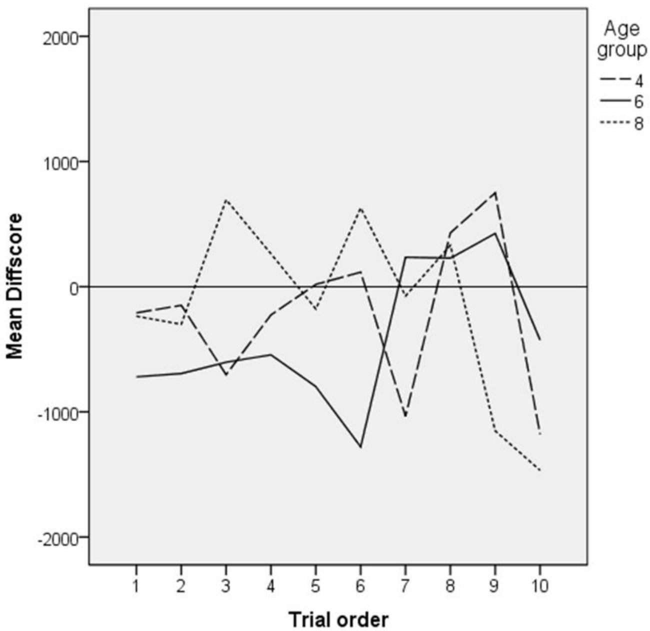


Figure 4.9: *The preference curves for all three Dutch age groups plotted as the difference between the mean TLTs during the SW and WS condition per trial pair.*

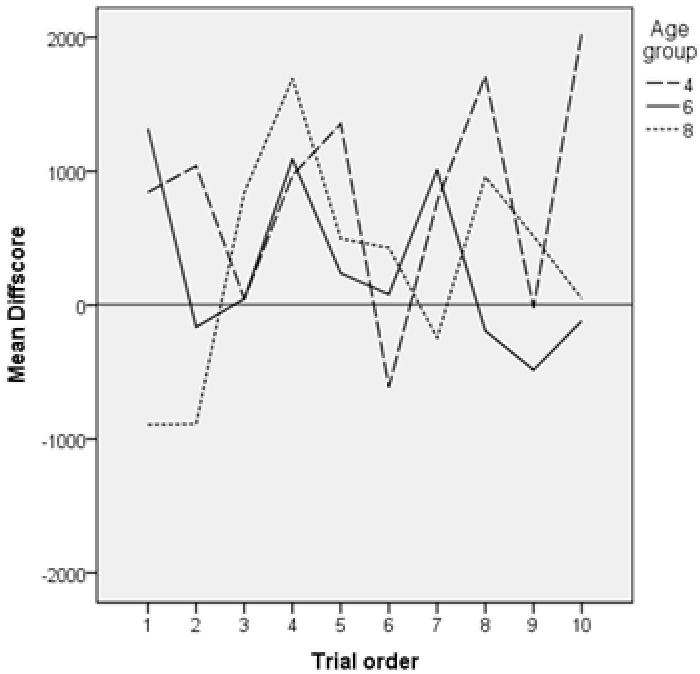


Figure 4.10: *The preference curves for all three Turkish age groups plotted as the difference between the mean TLTs during the SW and WS condition per trial pair.*

A Linear Mixed Model analysis was repeated, but this time with the normalized difference scores in ms as the dependent variable, language, age group and trial order as fixed factors, and participant as a random factor. I now find a significant main effect of language ($F(1, 191) = 20.499$, $p < .001$), which means the language groups differ in the direction of the difference scores, as the Dutch language group shows an overall SW familiarity preference, while the Turkish language group shows an overall SW novelty preference. As can be seen in figures 4.9 and 4.10, the preference development for the three age groups between the language groups seems to differ both in direction and timing. This is supported by a significant three-way interaction between the variables language, age group and trial order ($F(18, 1651) = 1.830$, $p = .018$). Consequently, I will discuss the Dutch and Turkish language groups separately.

In the Dutch language group, all three age groups start out with a slight familiarity preference, which is followed by an increase in familiarity preference for both the 4- and 6-month-olds, but an early shift to novelty for the 8-month-olds between trial pairs two and six. For the 6-month-olds, we can observe a similar shift to novelty much later in the experiment between trial pairs six and nine, and for the 4-month-olds the clearest shift to novelty is again somewhat later, namely between trial pairs seven and nine. When we look at the pairwise comparisons we find significant differences between the 6-month-old and both the 4- and 8-month-old age groups on trial pair six (p=.044 and p=.006, respectively). Here the 6-month-olds still show a familiarity preference, but the 8-month-olds already show a novelty preference. On trial pair nine, we find significant differences between the 8-month-olds and both 4- and 6-month-olds, (p=.012 and p=.022, respectively), for which the 8-month-olds seem to have shifted back to a familiarity preference, whereas the 4- and 6-month-olds have not (yet).

This shift back from novelty to familiarity is not predicted by the Hunter & Ames model, but it is what we see for all three age groups. However the shift is later for the 4- and 6-month-olds (between trial pairs nine and ten), who also demonstrated a later shift to novelty than the 8-month-olds. Even though this shift in preference back to the familiar stimulus is not predicted by the model, it does make sense from a habituation perspective. If infants have attended more to the novel stimulus for some time they will, after having encoded this stimulus, shift their attention back to the other less attended stimulus. Unfortunately, the Hunter & Ames model does not specify what happens after the shift to novelty has taken place. However, it should be noted that from trial pair five onwards there is data available from progressively less infants (from 102 infants in trial pair five to 82 infants in trial pair ten). In other words, the second half of the data might not be a random subsample considering that infants who make it to the end of the experiment might be more likely to shift to novelty later in the experiment or not to shift to novelty at all. Additionally, analysis over the second half of the experiment might be less reliable than over the first half of the experiment, simply because the former contains less data.

For the Dutch data we find the expected differences between the age groups in terms of the duration of the preferences stages. We find longer lasting familiarity preferences for the 4- and 6-month olds and an early shift to a novelty preference for the 8-month-olds. A rapid shift to novelty, according to Hunter & Ames, can be an indication of the task difficulty being too low: “If task difficulty were too low, familiarity preferences, although hypothetically possible, may appear so quickly after the onset of familiarization and for so brief a duration that, in practice, measuring them would be difficult.” This can provide an explanation for why we do find a familiarity preference for the stress pattern of the native language for the 6-month-old infants, but not for the 8-month-old infants.

In sum, we seem to find a preference for the SW stress pattern in the 4-month-old age group, but it is not (yet) sufficiently pronounced, as shown by the small effect size. This SW preference is consistent with the dominant native pattern, but need not necessarily come from the input of the native language. I will elaborate on the possibility of a universal trochaic bias, my third hypothesis, when discussing the Turkish data below. For the Dutch 6-month-old age group we do find a native preference for the SW stress pattern with a larger effect size. However, this preference is not retained in the 8-month-old age group, probably due to an early shift in preference to the novel WS stress pattern because of the low task difficulty for this age group.

In the Turkish language group the 4- and 6-month-old age groups seem to start out with a novelty preference for the SW stress pattern and the 4-month-olds keep this preference until the end of the experiment, without significant shifts towards the familiar WS stress pattern. The 6-month-olds show a shift towards familiarity between trial pairs seven and ten. Recall that according to Hunter & Ames the order in which the direction of the preferences takes place cannot be altered, in other words, familiarity preferences must appear before novelty preferences in order to interpret a novelty preference as such. As the SW preference in the 4- and 6-month-old Turkish-learning infants is not preceded by even a short WS preference, their preference for the SW stress pattern cannot be interpreted as a novelty preference. I therefore hypothesize that the results of the 4- and 6-month-old age groups are consistent with a universal trochaic bias and I will return to this in the conclusion section. However, in the 8-month-old Turkish group we see a pattern very similar to that of the 8-month-old Dutch group: an initial, but very short, familiarity preference, shifting to novelty already at trial pair two and shifting back to familiarity between trial pairs eight and ten. We also see a shift to the familiar WS pattern for the 6-month-old Turkish-learning infants at the end of the experiment, but not for the 4-month-olds.

When we look at the pairwise comparisons we indeed find significant differences between the Turkish-learning 8-month-olds and both 4- and 6-month-olds on trial pair one ($p=.006$ and $p=.004$, respectively), on which the 8-month-olds show a preference for the familiar WS stress pattern, whereas the 4- and 6-month-olds show a preference for the SW stress pattern. We still find a significant difference on trial pair two between the 8-month-olds and 4-month-olds ($p=.044$) whereby the 8-month-olds still show a familiarity preference and the 4-month-olds again show a preference for the SW stress pattern. On trial pair ten, we find significant differences between the 4-month-olds and both 6- and 8-month-old age groups ($p=.009$ and $p=.025$, respectively), as both the 6- and 8-month-olds show a shift in preference towards the familiar WS stress pattern, but the 4-month-olds do not. Again, a shift to a familiar stimulus after a novelty preference is not predicted by the model and we should be aware of the fact that also in the Turkish language group the second half of the experiment contains data from less infants than the first half of the experiment, decreasing from 92 infants in trial pair five to 69 infants in trial pair ten.

For the Turkish data, except for possibly the 8-month-old group, we do not find the expected differences in the duration of the preference stages between the age groups. The 4- and 6-month-old age groups do not show an initial preference for the familiar WS stress pattern, rather an immediate preference for the SW stress pattern. For the 4-month-olds this preference stays in place until the very end of the experiment, hence the larger effect size for the SW preference in the 4-month-old Turkish-learning group. However, in the 6-month-old age group there is a shift towards the familiar WS stress pattern at the end of the experiment, causing the smaller effect size for the SW preference in this age group. I hypothesize that the 4- and 6-month-old infants do not show a novelty preference, but rather a universal trochaic bias. I will further elaborate on this hypothesis in the conclusion section below. The Turkish-learning 8-month-olds show a very similar development of preferences during the experiment as the Dutch-learning 8-month-olds and we can therefore again conclude that the task difficulty of this experiment was probably too low for infants of this age. Nevertheless, the fact that the direction of preferences can and does change over time, thereby cancelling out strong preference effects in one direction, provides another explanation for the relatively small differences found between the conditions for each age group and in both language groups.

4.5 Conclusion

The results of the current experiment for the Dutch baseline group, starting with the 4-month-old infants, appear to be in line with the findings from the behavioral experiment by Höhle et al. (2009), where no preference for German-learning 4-month-olds was found. In contrast, the results do not replicate the findings from Friederici et al. (2007) who showed that German-learning infants already have some representation of the stress pattern of their native language at 4-5 months of age as demonstrated through the use of EEG measures. In my study, the Dutch-learning 4-month-olds were also tested in a behavioral paradigm. This explains the fact that they do not show a strong tendency for listening to the SW stress pattern of their native language. As for the 6-month-old Dutch-learning infants, I can also state that my results match those of earlier studies investigating the rhythmic preferences of German-learning infants. Results of Höhle et al. (2009) showed a preference in German-learning 6-month-olds to attend more to the native SW stress pattern than the non-native WS stress pattern. Therefore, the results from the Dutch-learning infants in this study demonstrating a clear preference for the native SW stress pattern at 6 months of age correspond well with the results from previous studies.

The results from the Dutch-learning 8-month-olds require a different interpretation. The current study was initially designed to test 4- and 6-month-old infants, but for comparability, I decided to run exactly the same experiment with the 8-month-old infants. In hindsight, however, we can conclude that this experiment was not appropriate for this age group. We do not find a clear preference for the SW stress pattern in the Dutch-learning 8-month-old group. This lack of a preference for the native stress pattern can be explained from the development of the preference curve in this group of infants during the experiment. Their preference development shows a very early shift to the novel WS stress pattern, indicating that the task difficulty of this experiment is too low for the 8-month-old infants.

To summarize, infants learning Dutch, which is very similar to German with respect to its dominant SW stress pattern, demonstrate a clear preference for the native SW pattern by 6 months of age, similar to German learning infants. The results of the Turkish-learning infants, acquiring a language with a dominant WS stress pattern, were needed to disentangle various accounts for the lack of a native language preference found in infants learning a WS dominant language thus far. I hypothesized that if Turkish-learning infants would not show any preference, as the French-learning infants in the study by Höhle et al. (2009), the lack of a rhythmic preference might be due to the syllable-timed character of these languages. However, if Turkish-learning infants would show a WS preference this would not only be evidence against a universal trochaic bias, but would also suggest that the SW preferences found thus far for English-, German- and Dutch-learning infants are indeed language-specific. If Turkish-learning infants, on the other hand, would also show an SW preference, this might be evidence for the existence of a universal trochaic bias, independent of the rhythm class of the language (stress-timed versus syllable-timed), but dependent on whether the language uses stress at the word level.

The results for the Turkish-learning infants provide evidence favourable to the last hypothesis. Contrary to previous studies with French-learning infants, we find an SW preference for all Turkish-learning age groups *combined* in this study. In previous studies testing French-learning infants, a language-specific response for the WS stress pattern was found at 4-5 months in an EEG experiment (Friederici et al., 2007), but no preference for either the SW or WS stress pattern was found at 6 months (Höhle et al., 2009) in a behavioral study, nor was there an effect of stress pattern in an earlier study with Turkish-learning infants (van Kampen et al., 2008). The current results of the Turkish-learning infants do show a preference for the SW stress pattern which is most pronounced at 4 months of age and much weaker at 6 and 8 months of age. These results could therefore be interpreted as evidence for the existence of an initial universal trochaic bias. However, if this interpretation were correct, we should, against earlier findings, also find a trochaic bias for French-learning 4-month-olds. Possibly, the trochaic bias hypothesis cannot explain all the data because a trochaic bias is not always found when testing young infants. One plausible explanation is that the trochaic bias is only 'activated' by the presence of word level stress in the language that is learned, which it is not the case in French.

I also consider a native-language dependency (NLD) hypothesis as an alternative explanation for the current findings. The NLD hypothesis states that infants will only demonstrate their knowledge of phonological patterns from their native language when presented with speech material of their native language. A more recent study by Segal & Kishon-Rabin (2012) has shown that Hebrew-learning 9-month-olds demonstrate a native WS preference when listening to a speaker of Hebrew, claiming that Hebrew has a dominant WS stress pattern similar to Turkish. However, when they presented another group of Hebrew-learning 9-month-olds with a speaker of English, a foreign language to them, the infants showed a non-native SW preference. Consequently, it seems to matter whether infants are listening to a speaker of their native language or to a speaker of a foreign language. Infants might resort to a universal trochaic bias when they do not recognize the stimuli they are listening to as coming from a speaker of their native language. This hypothesis can explain the results of Segal & Kishon-Rabin (2012) as well as my own results.

I would like to conclude with the following methodological point: both the Dutch- and Turkish-learning infants in my study were listening to a speaker of a foreign language, namely Spanish, and both of them show an SW preference. This preference might be independent of the language they are learning, as they are not listening to a speaker of their native language. This universal trochaic bias can, then, only be overwritten by input from a language with a different dominant stress pattern, as is demonstrated by the older Hebrew-learning infants presenting a native WS preference when listening to a speaker of Hebrew. The results of the Hebrew-learning infants listening to a speaker of English also imply that infants do not generalize an acquired stress pattern for their native language to a foreign language. This non-native speaker tendency is often the case for one or both of the language groups in cross-linguistic studies similar to my study. Therefore, adequate testing of language-specific rhythmic preferences may require a speaker of the native language. Consequently, I will present a follow-up study to test the NLD hypothesis where I will present infants with stimuli spoken by a native speaker of their language instead of a speaker of a foreign language. This follow-up experiment will be reported on in the next chapter, Chapter 5.

Chapter 5
The universal trochaic bias in monolingual Dutch- and Turkish-learning infants

5.1 Introduction

The previous chapter discussed an experiment conducted to test rhythmic preferences in the Dutch baseline group and showed an overall SW preference between 4 and 8 months of age. The results seemed to correspond with findings from the behavioral experiment by Höhle et al. (2009) testing German-learning infants with a similar goal and under similar test conditions. Akin to their results from German-learning 4-month-olds, I did not find a pronounced SW preference in Dutch-learning 4-month-olds. I did however find that Dutch-learning 6-month-olds prefer to listen to the native SW stress pattern over the non-native WS stress pattern, which is in line with the findings from Höhle and colleagues for German-learning 6-month-olds. Therefore the results of the 4- and 6-month old Dutch-learning infants resemble the results of the German study where a clear preference for the native SW stress pattern is only demonstrated at 6 months of age.

In regards to the Dutch-learning 8-month-olds, I concluded that the experiment did not suit this age group very well. Their preference development demonstrated a very early shift to the novel WS stress pattern, indicating that the task complexity of the experiment was too low for them. This may be due to the lack of segmental variation in the stimuli. Moreover, the experiment used fixed trials, thus these infants could not control the duration of the trials by their own looking behavior. These last two points regarding the lack of segmental variation and the use of fixed trials, apropos, apply to both groups of 8-month-old infants across the Dutch and Turkish language groups.

The goals of this thesis indicate that the preference results of the Turkish-learning infants – acquiring a language with a dominant WS word stress pattern – are needed to disentangle various accounts for why we have thus far failed to find a native word stress pattern preference in infants learning such a language. My results for the Turkish-learning infants reported in Chapter 4 appear to provide evidence for a Universal Trochaic Bias (UTB). Contrary to previous studies, I found an SW preference in the Turkish language group. The results of the Turkish-learning infants showed the clearest preference for the SW stress pattern at 4 months of age, but a much less clear SW preference at 6 and 8 months of age. The nature and time sequence of these results may be interpreted as evidence for a Universal Trochaic Bias.

This interpretation implies a general trochaic bias in early acquisition before 6 months of age, regardless of the dominant word stress pattern of the language. This, however, does not seem to be the case. Höhle et al. (2009) and Friederici et al. (2007) show no preference at all in German-learning infants at 4 months of age and a unidirectional language-specific discrimination of iambic patterns by French-learning infants at 4-5 months of age. Thus, this naturally raises the question of whether this range of results can be mutually reconciled. A number of possibilities come to mind. Firstly, it seems plausible to assume that the trochaic bias is only 'activated' by the presence of genuine word level stress in the language learned. There is no word level stress present in French (Grammont, 1965; Dupoux et al., 1997; Jun & Fougeron, 2000), whereas Turkish has word level stress. However, this assumption requires that infants first have to develop a lexicon in which phonological representations including word stress information are present, which is unlikely for infants below 6 months of age. Moreover, a lack of an initial trochaic bias is not limited to French-learning infants, as it is also not observed in German- and Dutch-learning infants (Höhle et al., 2009; this dissertation, Chapter 4) at 4 months of age.

Another explanation, which is more methodological in nature, may be that the infants in my study, and the French-learning infants in the studies by Höhle and colleagues (2009) and Friederici and colleagues (2007), were not tested with stimuli from a speaker of their native language, but from a speaker of a foreign language. A study by Segal & Kishon-Rabin (2012) showed that Hebrew-learning 9-month-olds demonstrate a native WS preference when listening to a speaker of Hebrew, as Hebrew has a dominant WS stress pattern at the word level, similar to Turkish. However, when they presented another group of Hebrew-learning 9-month-olds with a speaker of English, a foreign language to them, the infants showed a non-native SW preference. Consequently, in an experimental situation of the kind employed here it seems to matter whether infants listen to a speaker of their native language or of a foreign language. Infants may resort to a universal trochaic bias when they do not recognize the speaker they are listening to as a speaker of their native language. In light of this last consideration I proposed the native-language dependency (NLD) hypothesis at the end of Chapter 4. The NLD predicts that infants only demonstrate knowledge of, and preferences for, the specific sound patterns of their native language when presented with speech material resembling the phonetic properties of their native language. The scope of this hypothesis may be wider than just rhythmic preferences, extending to other prosodic properties such as tone and also segmental properties, but here I will focus on the suprasegmental domain of stress.

The NLD hypothesis implies that infants recognize their native language by, in this case, its prosodic characteristics by the age of 4 months. Bosch & Sebastián-Gallés (1997) demonstrate that native language recognition already occurs at the age of 4 months. In their experiments, monolingual Catalan- or Spanish-learning infants could distinguish between these two languages and, furthermore, recognize their native language (Catalan or Spanish) by only relying on suprasegmental information when confronted with low-pass filtered speech. Another study by Nazzi, Jusczyk & Johnson (2000) provides additional evidence for early native language recognition from monolingual American-English infants. By 5 months of age they could discriminate languages within the same rhythm class, as long as the native language was one of those presented, whilst earlier studies showed that newborns could only discriminate languages from different rhythm classes (Nazzi et al, 1998). Nazzi and colleagues (2000) thus label the native-language acquisition hypothesis as more likely than the rhythmic-class acquisition hypothesis. The native-language acquisition hypothesis states that infants acquire fine-grained acoustic-prosodic information about the rhythmic structure of their native language in the first few months of life. As the same rhythmic structure can have different phonetic correlates in different languages, these phonetic properties may form the basis for early native language recognition. Adequate testing of language-specific rhythmic preferences, therefore, seems to require a speaker of the native language, a condition that was not present in the experiment described in Chapter 4. For this reason, the decision was made to test the same language groups of Dutch- and Turkish-learning infants again, each with stimuli recorded from a speaker of their native language in a follow-up experiment.

5.2 Method

5.2.1 Participants

The first group of participants in this experiment are 24 Dutch-learning infants who were tested longitudinally at 6 and 8 months of age. This is different from the study described in the previous chapter in which the Dutch-learning infants were tested in a cross-sectional design. The reason for this is that the preference experiment described below was combined with a segmentation experiment at 8 months of age (see Chapter 7) in a longitudinal study.

The participants were only tested at 6 and 8 months and not at 4 months, because a native-language preference has thus far only been found, and is thus only expected, from 6 months of age. The mean age of the Dutch participants at 6 months was 200 days and at 8 months the mean age was 252 days. The participants were all healthy full-term infants, without any known hearing or visual impairments, and were being raised in a monolingual Dutch environment. Another four and three infants were tested at 6 and 8 months, respectively, but data from these infants could not be included due to crying, fussiness or technical issues.

The second group of participants in this experiment were 48 Turkish-learning infants: 24 infants around 6 months of age and 24 infants around 8 months of age, tested in a cross-sectional design. The intention was to follow the development and also test the Turkish group in a longitudinal design, but this proved to be impossible as the same caregivers were unfortunately not able to return to the lab when their infants reached the age of 8 months. The mean age of the Turkish 6-month-old age group was 197 days and of the 8-month-old age group the mean age was 256 days. The participants were all healthy full-term infants, without any known hearing or visual impairments, and were being raised in a monolingual Turkish environment. Another four and six infants were tested in the 6- and 8-month-old group, respectively, but data from these infants could not be included due to crying, fussiness, technical issues or parental interference. All participant information is summarized in Table 5.1.

Table 5.1: *Included and total number of Dutch- and Turkish-learning infants tested per age group with mean age and age range in days (longitudinal and cross-sectional design).*

Language group	*Dutch (longitudinal)*		*Turkish (cross-sectional)*	
Age group	6 months	8 months	6 months	8 months
Included (total)	24 (28)	24 (27)	24 (28)	24 (30)
Mean age	200	252	197	256
(range)	(174-218)	(237-274)	(174-221)	(236-282)

5.2.2 Material

The infants were presented with one of either two auditory stimuli, both disyllabic pseudo-words consisting of CVCCVC sequences: /mernel/ or

/darnam/. As the Dutch portion is a longitudinal study, it was important to have two segmentally different stimuli instead of the single pseudo-word used in the cross-sectional study described in Chapter 4. If infants were tested with /mernel/ at 6 months of age, they were tested with /darnam/ at 8 months of age and vice versa. The order was furthermore counter-balanced. Both of the pseudo-words are phonotactically well-formed in both Dutch and Turkish and either the first or the last syllable received stress, forming an SW or WS unit, respectively. The pseudo-words were, similarly to the pseudo-word used in the previous preference experiment described in Chapter 4, controlled for segmental and phonotactic naturalness by using phonemes and diphones with similar frequencies and transitional probabilities in Dutch and Turkish.

A balanced bilingual female speaker of Dutch and Turkish, who alternately lived in the Netherlands and Turkey throughout her life, recorded 12 tokens for each pseudo-word and each stress pattern in each language. The speaker was instructed to produce the pseudo-words in a Dutch carrier sentence (Ik ga … zeggen, "I go … say", 'I am going to say …') on one day and in a Turkish carrier sentence (Ben … dedim, "I … said", 'I said …') on another day, yet always in an infant-directed manner. The speaker had significant experience in working with infants in the baby lab of the Utrecht Institute of Linguistics OTS and therefore did not have difficulty speaking in an infant-directed manner. As presented in Table 5.2 for the Dutch version of the stimuli, stressed syllables (meaning the rhyme portion of the syllable as this is the domain of syllable duration) had an average duration of 197 milliseconds (ms), an average mean pitch of 403.8 Hz, and an average mean intensity of 70.0 dB, whereas unstressed syllables had an average duration of 183 ms, an average mean pitch of 261.8 Hz, and an average mean intensity of 68.2 dB. A Multivariate General Linear Model revealed that the stressed syllables (always only the rhyme portion) had a significantly longer duration (difference: 14 ms; $F(1,94) = 10.728$, $p=.001$), a significantly higher mean pitch (difference: 142 Hz; $F(1,94) = 449.107$, $p=.000$) and a significantly higher mean intensity (difference: 1.8 dB; $F(1,94) = 22.688$, $p=.000$) than (the rhyme portion of) the unstressed syllables.

For the Turkish version of the stimuli as presented in Table 5.3, stressed syllables (again the rhyme portion) had an average duration of 229 milliseconds (ms), an average mean pitch of 376.2 Hz, and an average mean intensity of 67.9 dB, whereas unstressed syllables had an average duration of 203 ms, an average mean pitch of 240.0 Hz, and an average mean intensity of 67.6 dB. A Multivariate General Linear Model revealed that the stressed syllables (again only the rhyme portion) had a significantly longer duration (difference: 26 ms; $F(1,94) = 31.381$, $p=.000$) and a significantly higher mean pitch (difference: 136.2 Hz; $F(1,94) = 520.956$, $p=.000$), but not a significantly higher mean intensity (difference: 0.2 dB; $F(1,94) = 0.252$, $p=.616$) than (the rhyme portion of) the unstressed syllables. As intensity, in general, is not a reliable cue to stress, the lack of a difference in intensity between stressed and unstressed syllables is not a concern.

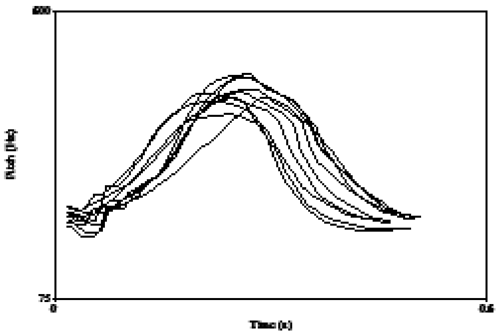
For both versions of the stimuli, for the Dutch as well as the Turkish version, pitch and duration were the most important phonetic cues to stress placement. However, this does not mean that the prosodic contour looked exactly the same in both versions of the stimuli. Figure 5.1 displays the different pitch contours for the different language versions of both SW and WS tokens of the stimuli. There are language-specific differences in the steepness of the pitch rise, it is steeper in the Dutch tokens than the Turkish. This conforms to what we know about Dutch and Turkish stress independently, as Turkish stress is reported to be marked by only a moderate pitch rise on stressed syllables (Konrot, 1981; Levi, 2005), whereas in Dutch, stress is marked by a large pitch rise on the stressed syllable when also receiving the phrasal accent (‘t Hart et al., 1990; Sluijter & van Heuven, 1996a; Sluijter et al., 1997). Due to the infant-directed manner in which the stimuli were pronounced, pitch movements are exaggerated for both languages as compared to adult-directed speech (ADS), but at the same time more attractive to infants and more capable of attracting and maintaining infants’ attention.

Table 5.2: *The average values over the 48 Dutch tokens for duration, mean pitch and mean intensity of the stressed and unstressed syllables with their standard deviation (SD).*

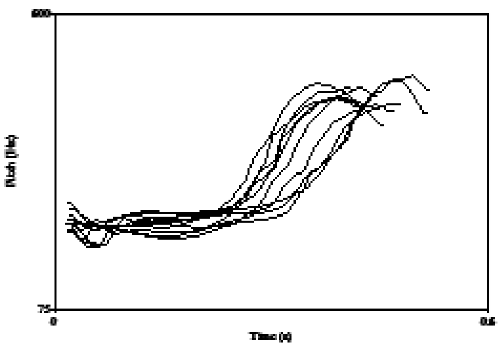
Syllable type	*Duration (ms)*	*Pitch (mHz)*	*Intensity (mdB)*
Stressed (SD)	197 (24.6)	403.8 (34.6)	70.0 (1.7)
Unstressed (SD)	183 (16.5)	261.8 (31.0)	68.2 (2.0)

Table 5.3: *The average values over the 48 Turkish tokens for duration, mean pitch and mean intensity of the stressed and unstressed syllables with their standard deviation (SD).*

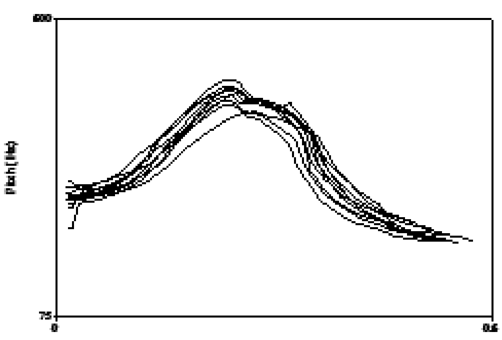
Syllable type	*Duration (ms)*	*Pitch (mHz)*	*Intensity (mdB)*
Stressed (SD)	229 (21.1)	376.2 (28.0)	67.9 (2.9)
Unstressed (SD)	203 (24.3)	240.0 (30.4)	67.6 (1.6)



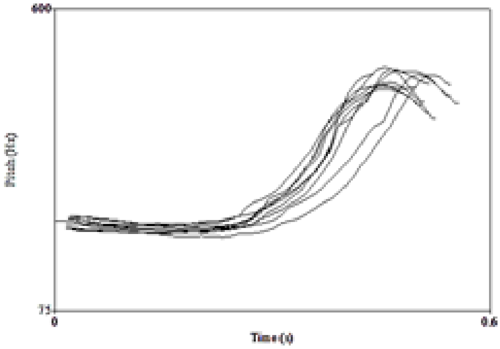
5.1a: *Dutch SW tokens of /darnam/*



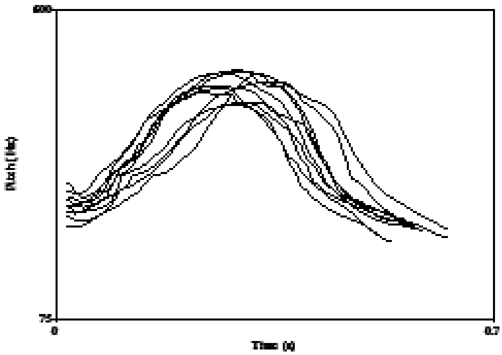
5.1b: *Dutch WS tokens of /darnam/*



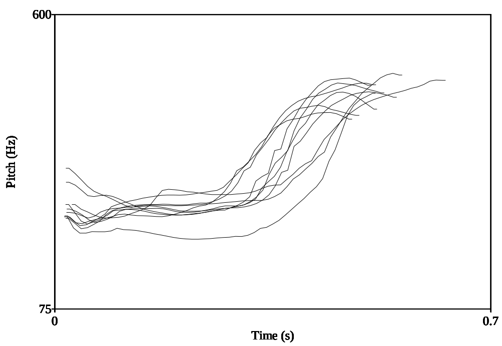
5.1c: *Turkish SW tokens of /darnam/*



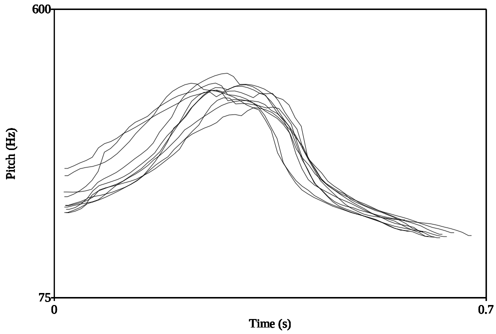
5.1d: *Turkish WS tokens of /darnam/*



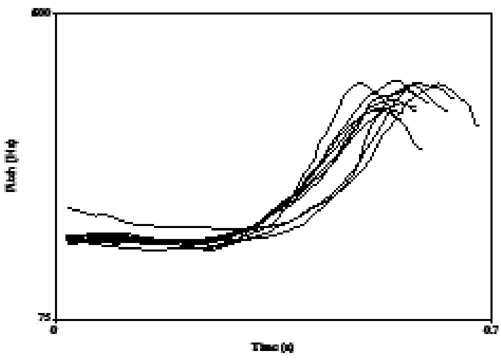
5.1e: *Dutch SW tokens of /mernel/*



5.1f: *Dutch WS tokens of /mernel/*



5.1g: *Turkish SW tokens of /mernel/*



5.1h: *Turkish WS tokens of /mernel/*

Figure 5.1a-h: *The pitch contours of all tokens (pitch in Hz plotted against time in seconds) for the different language versions (Dutch vs Turkish), the different stimuli (/darnam/ vs /mernel/) and the different stimuli-types (SW vs WS).*

In addition to the acoustical and statistical analysis, the stimuli were judged by adult native speakers of Dutch or Turkish in order to ensure that the correct language-specific stress cues were present in the recordings. The Dutch participants were 29 undergraduate students of Linguistics at Utrecht University and all native speakers of Dutch. The participants were asked to judge each stimulus on the position of the stressed syllable, that is, was the stress placed on the first or on the last syllable? The adult native speakers of Dutch judged the stress placement 99% correctly, meaning the participants indicated the syllable which possessed the phonetic cues to stress as the stressed syllable. In addition, they also listened to the Turkish version of the stimuli and were asked which set of stimuli sounded more native (more like Dutch) to them. The participants reported the Dutch version of the stimuli as sounding more native-like.

The Turkish participants for the stress placement assessment were two undergraduate students of Psychology at Koç University and both native speakers of Turkish. They judged the stress placement in 97% correctly and both reported the Turkish version of the stimuli to sound more native-like. Based on the native speaker judgments for both languages 20 tokens out of the 24 recorded tokens per stress pattern were selected, of which the stress placement was judged correctly by all native speakers. Out of these 20 tokens, 10 trials per stress condition containing 10 tokens with the same stress pattern and in a pseudo-randomized order, were created. The other 2 tokens per stress pattern were used to create a short trial of 4 seconds which was used to familiarize the infants with the auditory and visual stimuli. Within each trial there was an inter-stimulus interval (ISI) of 750 ms between each token, to prevent from an iterative pattern being perceived on more than two syllables.

Due to the fact that SW pseudo-words intrinsically had a longer duration than WS pseudo-words, the SW trials had a slightly longer duration than the WS trials. In the Dutch version of the experiment the SW trials had a mean duration of 12209 ms, ranging between 11747 and 12670 ms, whereas the WS trials had a mean duration of 11792 ms, ranging between 11507 and 12077 ms. In the Turkish version of the experiment the SW trials had a mean duration of 12675 ms, ranging between 12250 and 13100 ms, whereas the WS trials had a mean duration of 12492 ms, ranging between 12076 and 12908 ms. As this experiment uses infant-controlled trials based on gaze-contingency, the slightly different average duration of the trials per condition was not considered to be an issue.

5.2.3 Procedure

In the Dutch part of the study an EyeLink 1000 eye tracker (with a special infant illuminator) connected to a Coolermaster CPU Intel i7 960 3.2 GHz computer was used at Utrecht University in the Netherlands. The eye tracking device was placed inside a sound-attenuated room in which the caregiver was seated on a chair behind a table. The infant was seated in a special ‘Bumbo’ baby seat on the table in front of the caregiver which was aligned with the (moveable) monitor of the eye tracking device. The caregiver was instructed not to interact with the infant and listened to music over headphones during the experiment. The infant faced the 17” screen of the eye tracking device at a distance of approximately 65 cm (or 25.6 inches). Behind the screen two Tangent EVO E4 (20-100 Watt) loudspeakers were placed, through which the auditory stimuli were played at a comfortable loudness of 60-65 dB.

The procedure for the Turkish part of the study was largely the same as for the Dutch portion. The main adaptation was the use of a different eye tracker, namely a Tobii T120, connected to a Dell Precision T5500 computer at Koç University in Istanbul in Turkey. Both eye trackers are suitable for the experiments as they have been developed for use with infant participants. The eye tracking device was placed inside a sound-attenuated room at the university's Language and Communication Development Lab of the Psychology Department. In this lab the infant was seated in a car seat which was placed on the lap of the caregiver and aligned with the (moveable) monitor of the eye tracking device, again at a distance of approximately 65 cm (or 25.6 inches).

The experiment was programmed and run in ZEP-software (Veenker, 2014), version 1.6. The experiment started with an attention grabber - which is different from the one used in the original preference experiment described in Chapter 4 - showing a moving image of a girl accompanied by a chime sound and followed by a short familiarization trial. The visual stimulus during all of the test trials was the same static picture of the smiling female face as in the preference experiment described in Chapter 4 (Figure 5.2). Prior to each test trial the attention grabber appeared again in order to ensure the infant was looking at the screen. The test phase consisted of 20 alternating trials: 10 per stress condition, either starting with an SW trial or a WS trial (in counterbalanced lists). This follow-up experiment used infant-controlled trials, based on gaze-contingency, instead of fixed trials. The duration of each trial and, with that, of the entire experiment was thus dependent on the interest of the infant. More specifically, a trial ended when the infant looked away from the visual stimulus for more than 2 consecutive seconds. The fact that the infants can control the duration of the trials by their own looking behavior may reduce the chance of an early switch to a novelty preference in the 8-month-old infants, as was found in the results described in Chapter 4. Figure 5.2 illustrates the outline of the infant-controlled preference experiment using eye tracking.

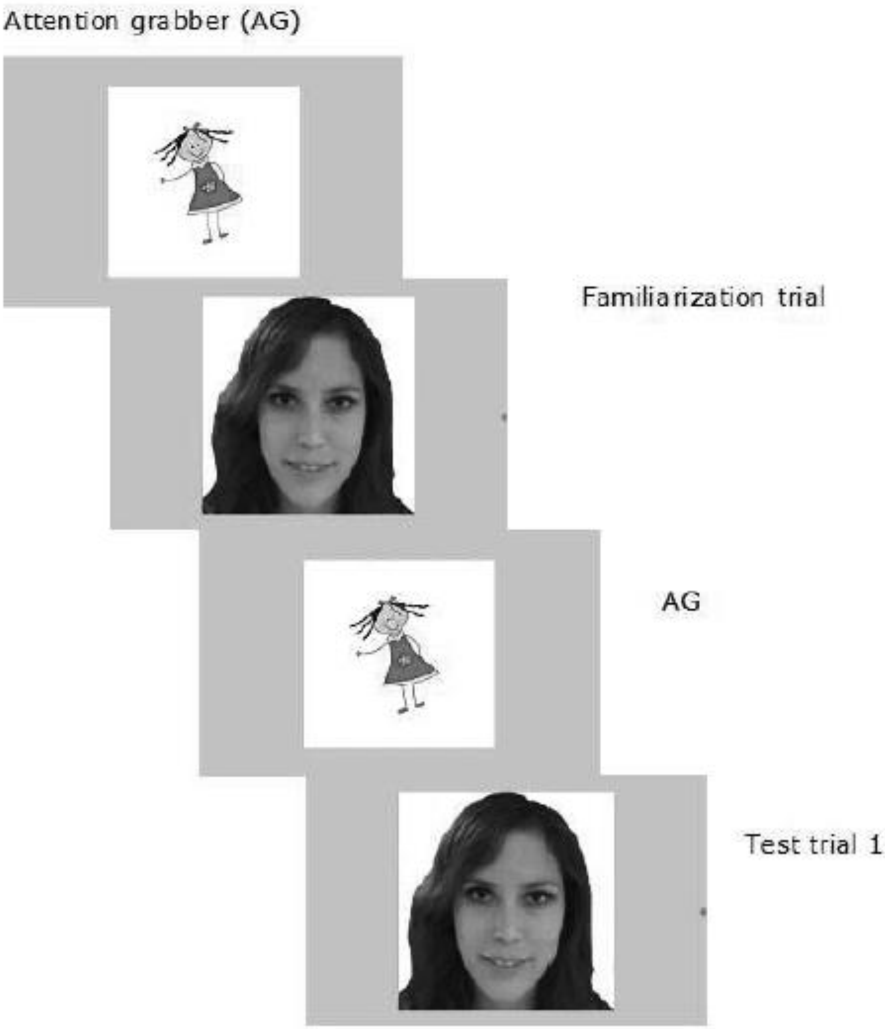


Figure 5.2: *The outline of the infant-controlled preference experiment using eye tracking.*

5.3 Data analysis

Several scripts were run on the raw data resulting from the ZEP-output in order to extract the information needed to calculate the Total Looking Time (TLT) per participant, per trial. First, the spatial and temporal criteria for 'looking' and 'looking away' were determined. The image of the female face was chosen as the Area of Interest (AoI), because this is the focus of the infants' attention when they are interested in the speech they are listening to simultaneously. The eye tracker records a gaze sample approximately every 4 or 8 ms (for the EyeLink 1000 and Tobii T120, respectively), providing the X- and Y-coordinates of the location of the eye gaze on the screen. In defining 'looking', I decided to not only include fixations (classified as the samples where gaze remained on approximately the same location on the screen for a minimum of 200 ms), but all gaze samples longer than 4 or 8 ms captured by the eye tracker on the AoI, which is also known as dwell time. In this way all looks to the AoI are counted towards the total looking time and no predefined criteria for what constitutes a 'look' is required. This also means that all gaze samples longer than 4 or 8 ms located outside the AoI were coded as 'looking away'.

If the experiment was stopped by the experimenter during a certain trial, this trial was excluded from the analysis entirely. The participant inclusion criteria stipulated that only infants who completed at least half of the experiment, that is, five or more trials per stress condition, were included in the analysis. This criterion serves to ensure the reliability of the sample, i.e. enough data points per infant and only of infants who in general paid enough attention to the experiment. This procedure follows one of the inclusion criteria used in Altvater-Mackensen, Van der Feest & Fikkert (2014).

5.4 Results

Recall that the purpose of this experiment is to see whether Dutch- or Turkish-learning infants at 6 and 8 months of age have a preference for their native language stress pattern when listening to a speaker of their native language. For both language groups combined, the mean TLT during the SW condition is 4392 ms (SD = 3548 ms) and the mean TLT during the WS condition is 4230 ms (SD = 3430 ms). These mean TLTs per condition for both language groups combined are presented in Figure 5.3. In order to determine whether there were significant effects of the variables condition, language group and age group, I analyzed the data by means of a Linear Mixed Model with condition, language group and age group as fixed factors, participant and trial as random factors and trial order as a covariate. I found a significant main effect of the variable condition ($F(1,1576) = 4.689$, $p=.031$), with a longer mean TLT during the SW trials than during the WS trials. This means that all infants, with both language groups and age groups combined, show a preference for listening to the SW stimuli over the WS stimuli. There is, however, no significant interaction of condition with either language or age group. This means the language and age groups behaved similarly with respect to the variable condition.

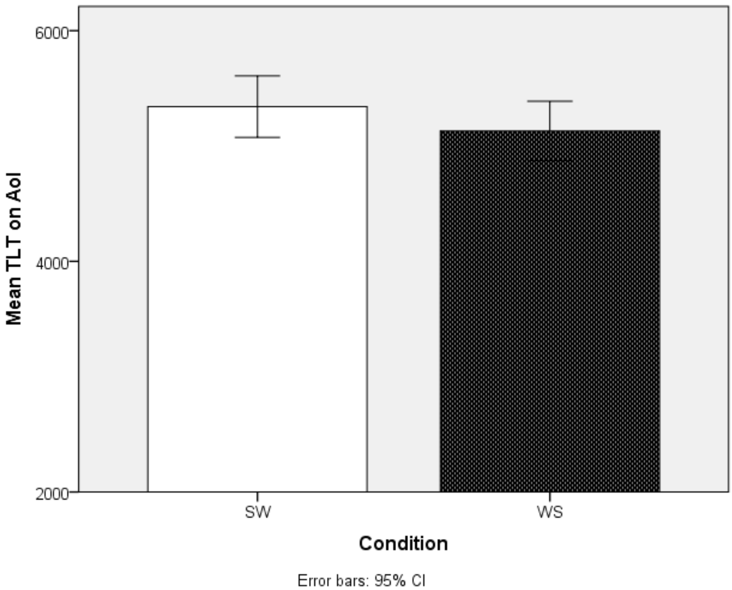


Figure 5.3: *The mean TLT on the AoI for all infants per condition (SW vs. WS) in ms.*

The results of the experiment as reported above indicate that we again cannot confirm a difference between the language groups, nor a development in rhythmic preference from 6 to 8 months of age. However, it is interesting to see what exactly is happening in each language group separately, as the difference in TLTs between the conditions seems to vary within each language group and per age group. The mean TLTs per condition for the Dutch language group per age group are shown in Figure 5.4 and Table 5.4. When we look at the effect sizes (ES in Cohen's d) for the factor condition within each age group, the ES is largest in the 6-month-old Dutch group (0.138), but much smaller and in the opposite direction in the 8-month-old age group (0.039). These results again suggest that for Dutch-learning infants the most pronounced SW preference occurs at 6 months of age. However, the Dutch-learning infants again do not seem to show a strong preference at 8 months of age, which is again unexpected as they seem to demonstrate a preference for the stress pattern of their native language at an earlier age. This suggests that adapting the experiment to an infant-controlled (gaze-contingent) experiment did not help in making the experiment more suitable for the 8-month-old infants. They show an increased interest for the novel WS pattern compared to the 6-month-old infants, similar to what the results from the experiment reported on in Chapter 4 displayed. Alternatively this could mean that the cognitive complexity is more dependent on the type of stimuli: I again used stimuli without segmental variation, while earlier studies have shown familiarity preferences at 9 months with segmentally varied stimuli (Jusczyk et al., 1993a; 1993b).

In sum, I replicated the results from the experiment in Chapter 4 for the Dutch-learning infants. They behave similarly when listening to a foreign language speaker and a native language speaker. However, in order to rule out that the trochaic bias I found is dependent on listening to a foreign language speaker, we first have to examine the behavior of the Turkish-learning infants in this follow-up experiment.

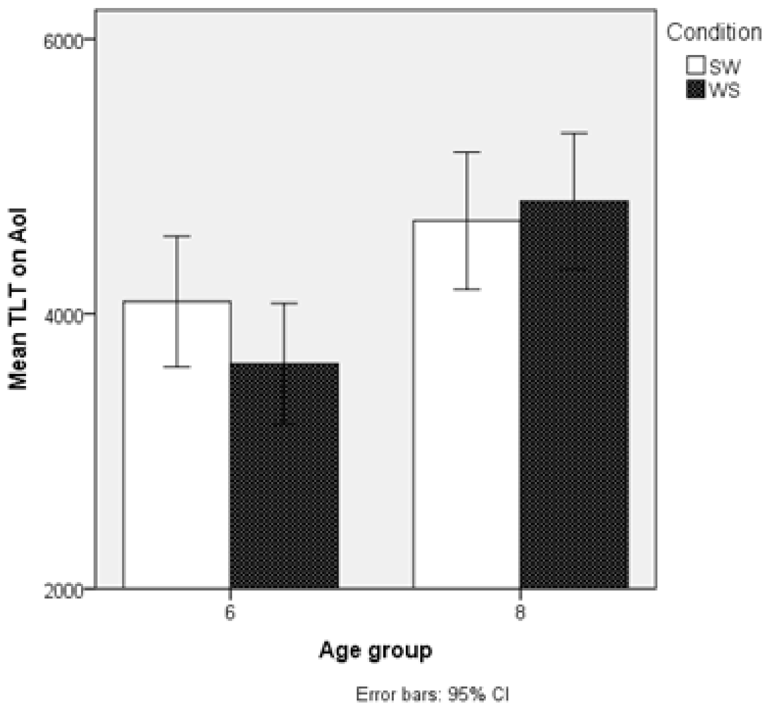


Figure 5.4: *The mean TLT in ms per Dutch-learning age group per condition (SW vs. WS).*

Table 5.4: *The mean TLT in ms per Dutch-learning age group per condition (SW vs. WS).*

Stress pattern	*6-month-olds*	*8-month-olds*
SW condition (SD)	4089 (3398)	4677 (3670)
WS condition (SD)	3636 (3170)	4817 (3582)

The mean TLTs for the Turkish language group per age group are presented in Figure 5.5 and Table 5.5. When we look at the effect sizes (ES in Cohen's d) for the factor condition within each age group, both the ES in the 6-month-old Turkish group and in the 8-month-old Turkish group are very small (0.060 and 0.075, respectively). These results again suggest that the Turkish-learning infants do not show a pronounced SW preference at either 6 or 8 month of age. Thus, I have also replicated the results from the original experiment described in Chapter 4 for Turkish-learning infants in these particular age groups. They behave similarly when listening to a foreign language speaker and a native language speaker. In sum, I replicated the results from the experiment in Chapter 4 overall. Both Dutch- and Turkish-learning infants behave similarly when listening to a foreign language speaker and a native language speaker. In other words, the NLD hypothesis does not seem to be confirmed, at least not for these age groups and languages groups.

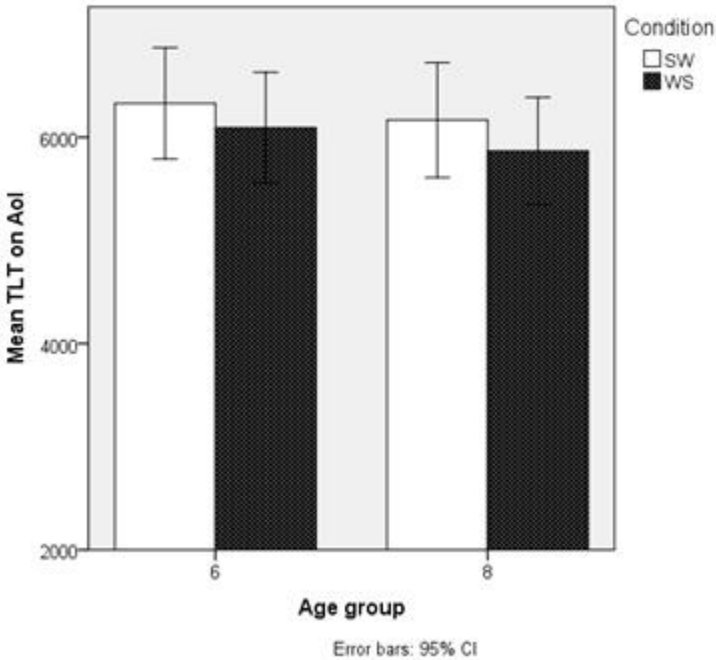


Figure 5.5: *The mean TLT in ms per Turkish-learning age group per condition (SW vs. WS).*

Table 5.5: *The mean TLT in ms per Turkish-learning age group per condition (SW vs. WS).*

Stress pattern	*6-month-olds*	*8-month-olds*
SW condition (SD)	6329 *(3996)*	6167 *(4152)*
WS condition (SD)	6092 *(3968)*	5867 *(3878)*

5.5 Discussion and conclusion

The aim of the follow-up experiment discussed here was to clarify the nature of the overall trochaic bias found in monolingual Dutch- and Turkish-learning infants, relative to the study described in Chapter 4. On the one hand I hypothesized that the results may be due to the fact that both Dutch and Turkish possess word level stress, whereas French does not. The idea would be that a universal trochaic bias is only 'triggered' in infants who are exposed to native input from a language with word level stress. However, before we are able to draw final conclusions from the original preference experiment, we need to ascertain that the explanation is not to be found in the phonetic cues of the stimuli used. My alternative hypothesis stipulated that the overall trochaic bias might have been induced by the fact that neither language group was listening to stimuli spoken by a speaker of their native language during the experiment. As predicted by the Native-language dependency (NLD) hypothesis, infants may resort to a universal trochaic bias when they do not recognize the language they are listening to as their native language, as a consequence of non-native phonetic cues.

We tested the NLD hypothesis by recording two versions of the stimuli spoken by a balanced bilingual Dutch-Turkish speaker and by presenting the Dutch version of the stimuli to the Dutch-learning infants and the Turkish version of the stimuli to the Turkish-learning infants. If the overall trochaic bias found in the study described in Chapter 4 was indeed induced by the fact that the infants listened to a speaker of a foreign language, I expected to find native preferences for both language groups in this follow-up experiment. More precisely I expected Dutch-learning infants to show a native SW preference from 6 months of age and Turkish-learning infants to present a native WS preference from 6 months of age. However, if no native preferences for either language group were found, this would not support the NLD hypothesis and the results from both the original experiment and the follow-up experiment would require an alternative explanation.

The results from the follow-up experiment for both language and age groups combined suggest an overall SW preference similar to the overall SW preference found in the original experiment. This combined result indicates that there is no effect of changing the speaker of the stimuli to a native speaker for either language group. When we look at the language and age groups separately we find a pattern of results which is similar to the pattern found in the original experiment. The Dutch-learning 6-month-olds again show the clearest SW preference, suggesting they demonstrate an SW preference independent of whether they are listening to a speaker of a foreign language or a speaker of their native language. The Dutch-learning 8-month-olds once more do not present a clear preference for either SW or WS stimuli, but again show an increased interest in the WS stimuli compared to the 6-month-old Dutch-learning infants. This additionally indicates that replacing the fixed trials from the original experiment with infant-controlled trials in the follow-up experiment did not prevent the 8-month-old Dutch-learning infants from developing an increased novelty preference for the WS stimuli. For this reason, it seems plausible that it is the lack of variation in the stimuli which poses too little of a challenge to the 8-month-old infants: stimuli without segmental variation are too low in cognitive complexity for this age group and therefore, finding a familiarity preference for this age group using this type of stimuli cannot to be expected.

The Turkish-learning infants do not show a clear preference for either SW or WS stimuli in either age group in this follow-up experiment. In the original experiment a clear preference for SW stimuli was only found in the 4-month-old age group, but not in the 6- and 8-month-old infants. Thus, there is also no effect of changing the speaker of the stimuli from a speaker of a foreign language to a speaker of their native language for the Turkish-learning infants. This result does not support the NLD hypothesis either. The results from both the original experiment and the follow-up experiment require an alternative explanation due to the lack of a trochaic bias in Dutch- and German-learning 4-month-olds. In the next chapter, Chapter 6, I will discuss the joined results from Chapters 4 and 5 against the background of different hypotheses which were introduced in the Literature review (Chapter 2).

Chapter 6
Interim discussion

In the Literature review and in Chapters 4 and 5, different hypotheses for infants' rhythmic preferences were introduced and tested in the experiments described in these chapters. In this interim discussion I will review the hypotheses once more and discuss what conclusions we can draw from the results of the two preference experiments. This interim discussion is also essential to the predictions that will be made regarding the use of either universal or language-specific stress cues for word segmentation by Dutch- and Turkish-learning infants at 8 months of age, which will be the topic of the next chapter, Chapter 7.

I will first discuss the Rhythm Class Hypothesis (RCH) and the rhythmic-activation proposal (Nazzi et al., 1998; 2006), then the Word level Stress (WLS) hypothesis (Höhle et al., 2009), the Universal Trochaic Bias (UTB) hypothesis (Allen & Hawkins, 1978; 1980), the native-language dependency (NLD) hypothesis (Chapters 4 and 5, this dissertation), and end with a discussion of the Iambic-Trochaic Law (ITL) and phrasal prominence proposal (Nespor et al., 2008). I will discuss all of these hypotheses critically and conclude that they are essentially insufficient in explaining the existing data. I will further argue that it is necessary to formulate separate predictions about word- and phrase level stress and also that these two types of stress play different roles at different developmental stages.

6.1 RCH and rhythmic-activation proposal

The Rhythm Class Hypothesis (RCH: Nazzi et al., 1998) in combination with the rhythmic-activation proposal (Nazzi et al., 2006) predicts that infants are sensitive to rhythmic differences between languages early on (i.a. Mehler et al., 1988; Nazzi et al., 2000) and that only infants learning a stress-timed language will demonstrate native rhythmic preferences. The results from the 6-month-olds in both of the preference experiments provide support for this hypothesis: infants learning Dutch – a stress-timed language – show a native SW preference, similar to German-learning infants at this age (Höhle et al., 2009), and infants learning Turkish – a syllable-timed language – do not show a clear preference (Chapter 5, this dissertation), similar to French-learning infants at this age (Höhle et al., 2009). Nevertheless, this combined hypothesis does not explain why Dutch-learning infants at 4 months of age do not show a clear SW preference, as they are able to recognize the stress-timed properties of their native language from birth (Mehler et al., 1988), and why Turkish-learning infants of the same age do show a clear SW preference. To comply with the RCH, the results from both of these language groups at 4 months thus have to be accounted for otherwise, for instance, by looking at other factors which may play a role in this age group. I will come back to these factors further on in this interim discussion.

6.2 Word level Stress hypothesis

To account for the lack of a rhythmic preference in French-learning 6-month-olds, Höhle and colleagues (2009) propose the WLS hypothesis. This hypothesis predicts that only infants learning a language which has stress at the word level, regardless of whether the language is stress-timed or syllable-timed, will learn to represent stress at the phonological level. This may explain why French-learning 6-month-olds do not show a preference for a particular stress pattern (Höhle et al., 2009), as French does not have stress at the word level (Dupoux et al., 1997). Additional evidence for this account comes from the fact that infants learning Spanish, also a syllable-timed language – but with lexical stress – do represent stress at the phonological level by 9 months of age (Skoruppa et al., 2009).

The results found in the two preference experiments described in this dissertation do not provide support for this hypothesis. Turkish infants do not show a preference for the stress pattern of their native language by 6 months of age despite learning a language with word level stress according to Inkelas (1999), Inkelas & Orgun (2003) and Göksel & Kerslake (2005) (cf. Levi, 2005), and following Kabak & Vogel's (2001) analysis of the Phonological Word (PW) in Turkish. The reason for this may be that Turkish, although marking stress at the word level, would not be classified as having lexical stress as most words in Turkish have default final stress with only a limited set of exceptions (Kabak & Vogel, 2001) and phonemic stress is rare, in contrast to Spanish in which stress position is much more variable and has to be lexically specified. In this respect Dutch is more similar to Spanish, as Dutch stress is phonemic at the word level. However, the WLS hypothesis, again cannot explain why Turkish-learning infants do demonstrate an SW preference at 4 months of age, as a rhythmic preference would not be expected at any age for Turkish-learning infants under this hypothesis. Therefore, something other than the mere notion of word level stress is needed to explain this early SW preference in Turkish-learning 4-month-old infants.

6.3 Universal Trochaic Bias hypothesis

As first proposed by Allen & Hawkins (1978; 1980), the UTB hypothesis predicts that all infants, regardless of the language they are learning, would initially prefer SW units over WS units, as this is claimed to be an innate human perceptual bias. Most evidence for this hypothesis is provided by truncation patterns in the early word productions of toddlers, both from stress-timed trochaic languages, such as Dutch (Fikkert, 1994), and languages with a word level iambic pattern, such as Hebrew (Adam & Bat-El, 2008; 2009). Early word productions have proven to provide a good insight into early word form representations at the segmental level in infants between 14 and 18 months of age (Levelt, 2012). Levelt (2012) provides evidence for the idea that phonological processes in production reflect infants' speech perception and, with that, their current phonological representations. The universal, thus language-independent, trochaic tendency in early speech production would then predict that all infants, regardless of the dominant stress pattern of the language they are learning, would initially prefer SW over WS patterns in perception.

The results from the Turkish-learning 4-month-olds in the experiment discussed in Chapter 4 seem to provide evidence for this initial trochaic bias because they show a clear preference for SW stimuli over WS stimuli at this age. The existence of a universal bias could only be confirmed, however, if we would also find clear SW preferences in Dutch-, German- and French-learning infants at the same age, which is not the case (Chapter 4, this dissertation; Höhle et al., 2009). Ultimately, an alternative account which explains the existing data more closely needs to be considered as the UTB hypothesis cannot explain all data gathered thus far, that is that there are initial trochaic biases in production or perception for some languages (Hebrew and Turkish, respectively) but for other languages only in production and not in perception (Dutch and German), and neither in production nor in perception (French).

In Chapters 4 and 5, a hybrid form of the WLS and UTB hypothesis was also considered. The hybrid hypothesis states that a universal trochaic bias is only 'activated' in infants who are exposed to a language with word level stress. This hypothesis could explain the lack of a rhythmic preference in 6-month-old French-learning infants and the consequent finding of a trochaic preference for infants learning languages with word level stress, such as English, German and Dutch, at 6 months of age. This hybrid hypothesis can also explain the initial SW preference of the Turkish-learning 4-month-olds reported on in Chapter 4, but it does not possess much explanatory power in predicting native rhythmic preferences for different languages groups and older age groups.

6.4 Native-language dependency hypothesis

Another explanation for the SW preference of the Turkish-learning 4-month-olds was the Native-language dependency (NLD) hypothesis proposed in Chapter 4 and tested with the experiment reported on in Chapter 5. This hypothesis states that only if infants are listening to a speaker of their native language, can native preferences be elicited, but if infants are listening to a speaker of a foreign language, they will demonstrate a ‘default' universal bias. This hypothesis was based on results obtained by Segal & Kishon-Rabin (2012) who investigated Hebrew-learning 9-month-olds with both a speaker of Hebrew and a speaker of English, a foreign language to the infants. They found that the infants showed a native WS preference when listening to a speaker of their native language, whilst showing an SW preference when listening to a speaker of English. Although the typological distance between Turkish and Spanish is about the same as between Hebrew and English, my experiments cannot provide evidence for this hypothesis as the Turkish-learning 6- and 8-month-olds reported on in Chapter 5 do not present a native WS preference when listening to a speaker of their native language instead of a speaker of a foreign language, Spanish in this case. A way of interpreting the Hebrew results differently, and simultaneously maintaining the current findings, may be as follows.

Regarding the phonetic cues for stress present in the Hebrew stimuli from Segal & Kishon-Rabin (2012), the WS stimuli clearly demonstrate a difference in duration between the stressed and unstressed syllables combined with only a slight difference in pitch, according to the authors. These WS stimuli are pitted against SW stimuli which do not possess a clear durational difference, but also only possess a slight pitch difference between the stressed and unstressed syllables. Yet, in the English stimuli from Segal & Kishon-Rabin (2012), the WS stimuli also present a clear difference in duration and a slight difference in pitch, and are pitted against SW stimuli without a clear durational difference, but more importantly, a considerable difference in pitch. As infants at least until 8 months of age are, as far as we know more sensitive to pitch information than to durational cues (Bion, Benavides-Varela & Nespor, 2011) it may be that the SW stimuli were simply more attractive to them than the WS stimuli when listening to the English speaker, due to the large pitch difference between the syllables in the SW stimuli. Bion and colleagues (2011) also claim that durational sensitivity may be more dependent on linguistic experience and this can, in my view, explain why a native WS preference was found when Hebrew-learning infants are listening to a speaker of a language with which they actually have experience, their native language.

To summarize, the NLD hypothesis does not correctly predict, nor sufficiently explain the data found in the follow-up experiment reported in Chapter 5. The results of this follow-up experiment do replicate the results from the initial experiment described in Chapter 4 for the 6- and 8-month-old infants of both the Dutch and Turkish language groups and, with that, confirm the validity of the results. Now that the consistency of results has been demonstrated, they must be explained in a different way. An early sensitivity to pitch, as already mentioned above, but also distinguishing between word- and phrase level stress, as will be explained in the next section, may form key elements in explaining the results from the current study, as well as in correctly predicting results of future studies with similar aims.

6.5 Combined ITL and phrasal prominence proposal

The differential processing of various kinds of phonetic cues, such as duration, pitch and intensity, by humans is captured in the Iambic-Trochaic Law (ITL: Woodrow, 1909; 1911; 1951 and Hayes, 1985; 1987; 1995). This law states that sequences of elements differing in pitch or intensity are processed as prominence-initial, in other words trochaically, whereas sequences of elements differing in duration are processed as prominence-final, or iambically. This law has been found not to be specific to language, but also holds for the musical (e.g. tones) and the visual domain (Hay & Diehl, 2007; Kurby & Zacks, 2008; Swallow & Zacks, 2008; Swallow, Zacks & Abrams, 2009). Within the language domain it means that the acoustic cues present in linguistic stimuli can influence the preference that infants demonstrate towards SW and WS stress patterns. If stress is mostly realized by pitch or intensity, the ITL predicts that infants will show an SW preference, whereas if stress is mostly realized by duration, infants would accordingly prefer WS patterns. Note that sensitivity to pitch is present early on, whereas sensitivity to duration may have to be acquired by infants (Bion et al., 2011). In the stimuli used in the two preference experiments described in Chapters 4 and 5 all three cues were present: the stressed syllables always had a longer duration and a higher pitch and intensity than the unstressed syllables. However, if young infants, until 8 months of age, are initially more sensitive to pitch cues than to durational cues, as processing duration information may be dependent on linguistic experience (Bion et al., 2011) and consequently, prefer SW groupings, this may have induced the overall trochaic bias found in both experiments for both language groups.

Nevertheless, there are interesting differences between the age groups in both language groups. These differences could be explained by approaching the ITL from a different perspective, that of the phrasal prominence proposal by Nespor et al. (2008). Departing from the idea that at the level of the phonological phrase languages can have either initial stress (on the first content word of the phrase) or final stress (on the last content word of the phrase), Nespor and colleagues (2008) proposed that, in accordance with the ITL, phrase initial stress is marked by increased pitch and phrase final stress by increased duration. Furthermore they suggest that these patterns are directly related to the word order within phrases, which also differs between languages (Nespor, Guasti & Christophe, 1996). When phonological phrases in a language have a complement-head order (object-verb), this is reflected in phrase initial stress marked by increased pitch. On the other hand, when a language has a head-complement order (verb-object), this is reflected in phrase final stress marked by increased duration.

Nespor and colleagues (2008) demonstrate with measurements of adult-directed speech that languages, such as German, which can have both word orders depending on clause type, mark these differentially by using a final increase in duration for verb-object (VO) main clauses and an initial increase in pitch for object-verb (OV) subordinate clauses. Dutch, which is very similar to German in this respect, also has both word orders, and I assume both phrasal prominence patterns occur in the linguistic input to infants learning this language as well. Turkish, however, is categorized as an OV language, although it has a relatively free word-order compared to German and Dutch, and, thus, an initial prominence pattern at the phrase level is dominant in the linguistic input to Turkish-learning infants, as predicted by measures of (adult-directed) Turkish reported on by Nespor et al. (2008), but see Kabak & Vogel (2001), Kabak & Revithiadou (2009a) and Kabak & Revithiadou (2009b) for a similar analysis of Turkish.

Considering the fact that infants are already sensitive to these different phrasal prominence patterns between 1 and 4 months of age (Christophe, Nespor, Guasti & Van Ooyen, 2003), but only start acquiring their first phonological word representations by 6 months of age (Bergelson & Swingley, 2012), it is plausible that the 4-month-old infants tested in the preference experiment described in Chapter 4 actually do not distinguish between the word and the phrase level in processing, whereas the infants from 6 months of age onward tested in both preference experiments reported on in Chapters 4 and 5 are starting to distinguish between the word and the phrase level when processing speech stimuli.

The idea that processing can take place at different levels, depending on developmental stage has strong implications for the expected preferences. For Dutch-learning 4-month-olds I would not expect to find any strong preference, as both phrasal prominence patterns, due to the occurrence of both word orders (SVO and SOV), are present in this language, whereas for Turkish-learning infants at this age I would expect a preference for an initial prominence pattern, based on the idea that infants are still phrase level processors before 6 months of age. These predictions are indeed confirmed by the Dutch and Turkish data from the experiment described in Chapter 4: Dutch-learning 4-month-olds do not show a strong preference for either pattern, whereas Turkish-learning 4-month-olds show a clear preference for the initial prominence (SW) pattern. Additional evidence for this idea can be found in data from German-learning infants, also not demonstrating a preference for either pattern at 4 months of age (Höhle et al., 2009). Furthermore, data from French-learning infants showing language-specific discrimination of the WS pattern at 4 months of age (Friederici et al., 2007), which may represent the final prominence pattern of French at the phrase level (Dupoux et al., 1997), also confirm this idea.

An interesting test case for the phrasal prominence proposal would be to test English-learning infants on their rhythmic preferences at 4 months of age, which to my knowledge has not yet been done or, at least, not been published. The prediction according to the phrasal prominence proposal would be that English-learning infants would actually show a final prominence (WS) preference at this age, because English has VO (head-compliment) word order. It has been shown that when English-learning 7-to-8-month-olds are confronted with a continuous sequence of non-linguistic tones (without any phrasal boundary marking), they demonstrate an WS pattern as their perceptual grouping bias and this effect is claimed to be caused by the WS phrasal prominence pattern of English (Yoshida, Iversen, Patel, Mazuka, Nito, Gervain & Werker, 2010).

For infants at 6 months of age the predictions are slightly more complicated, as infants at this age are still 'beginners' at word level processing. For Dutch, the prediction would be that as both patterns occur at the phrase level, but the initial prominence (SW) pattern dominates at the word level, this pattern receives a frequency 'boost' in the linguistic input infants process and therefore, Dutch-learning infants are expected to show an SW preference at 6 months of age. This prediction is confirmed by the data from both preference experiments reported on: Dutch-learning infants demonstrate a clear SW preference at 6 months of age. For Turkish the prediction would be that as one pattern (SW) is dominant at the phrase level, but the opposite pattern (WS) is dominant at the word level, these conflicting patterns do not cause one of the patterns to be more dominant in the linguistic input infants process and, therefore, Turkish-learning infants are not expected to show a preference for either pattern at 6 months of age. This prediction is also reflected in the data from both preference experiments described in the previous chapters: Turkish-learning infants do not demonstrate a clear preference at 6 months of age.

Again, additional evidence for this idea can be found in previous studies examining infants learning German, which also receives this frequency 'boost' of the SW pattern at the start of word level processing and German-learning infants, similarly to Dutch-learning infants, show an SW preference at 6 months of age (Höhle et al., 2009). Studies investigating infants learning English, in which there is a potential conflict between the phrase level pattern (WS) and the word level pattern (SW), also do not show a clear preference yet at 6 months of age, but only at 9 months of age (Jusczyk et al. 1993a; Morgan, 1996). It may therefore be the case that Turkish-learning infants will show a WS preference at the age of 9 months, similar to Hebrew-learning 9-month-olds demonstrating a native word level WS preference when listening to Hebrew (Segal & Kishon-Rabin, 2012). This would need to be tested by presenting Turkish-learning 9-month-olds with segmentally varied SW and WS stimuli spoken by a native speaker.

In regards to the 8-month-olds tested in the two preference experiments reported on here, the analyses primarily indicate that the experiments were not suitable for 8-month-olds. Most studies testing infants around this age use segmentally varied stimuli and do find the predicted results (Jusczyk et al. 1993a; Morgan, 1996; Segal & Kishon-Rabin, 2012). It seems plausible that the task complexity of this design, using stimuli without segmental variation, is simply insufficiently challenging cognitively for this age group. Evidence for this comes from the very early shifts to novelty preferences in both the Dutch- and Turkish-learning 8-month-old groups, comprising the possibility of finding native familiarity preferences at this age. Consequently, 8-month-old infants should, at least, be tested with segmentally varied stimuli to be able to find native preferences. I would in that case expect to find a native SW preference for 8-month-old Dutch-learning infants, although for Turkish-learning infants tested at 8 months of age with segmentally varied stimuli I am not certain whether a native WS preference can already be expected, due to having to resolve the conflict between phrase- and word level stress patterns.

To conclude, a combination of the ITL and the phrasal prominence proposal may account for the preferences found in the age and language groups tested in this study. At 4 months of age, both Dutch- and Turkish-learning infants' rhythmic preferences reflect the occurrence of phrasal prominence patterns in those languages. At 6 months of age, the rhythmic preferences of both Dutch- and Turkish-learning infants reflect the increasing dominance of the word level stress pattern as they start building their initial word representations. At 8 months of age, neither Dutch- nor Turkish-learning infants show a rhythmic preference in either of the experiments reported on in Chapter 4 and 5, most probably reflecting the unsuitability of the current experimental design for this age group.

In future research this differential processing of word and phrase level stress by infants at different developmental stages and ages, i.e. prelexical and lexical infants, should be further examined. This can be conducted by making the distinction between stimuli with and without phrasal boundaries more clear in experimental designs. When infants are presented with speech stimuli without phrasal boundaries, it may induce segmentation based on phrase level stress (e.g. Yoshida et al., 2010) whilst stimuli containing clear phrasal boundaries may induce segmentation based on word level stress. These two levels of processing cannot be distinguished when infants are presented with lists of isolated words, as an isolated word also constitutes a phrase. The implication this interim discussion has for the predictions made regarding Dutch- and Turkish-learning infants' use of stress cues for speech segmentation at 8 months of age will be elaborated upon in the next chapter, Chapter 7.

Chapter 7
The use of stress cues for word segmentation by monolingual Dutch- and Turkish-learning infants

7.1 Introduction

In this chapter, I take a different perspective from the approach discussed previously, although the previous results are gratefully taken into account. In Chapters 4 to 6, I examined whether infants prefer to listen to the dominant stress pattern of their native language and if so, at which age they do so. It was found that Dutch-learning 6-month-olds demonstrate a preference for the dominant word level stress pattern of Dutch, but not yet at 4 months of age. At the same ages, Turkish-learning 4-month-olds demonstrate a preference for the dominant phrase level stress pattern of Turkish; however, no preference for the native word level stress pattern was found at 6 months of age. Recall that the main aim of this dissertation is to examine if infants use rhythmic cues to segment the speech stream into word-like units. As both Dutch- and Turkish-learning infants have been shown to be sensitive to (different) rhythmic patterns related to their native language in the preference experiments, I expect infants to use these patterns in their speech segmentation strategies as well. Therefore, in this chapter, I change perspectives towards examining the role of stress cues in speech segmentation, in particular, in word segmentation. I will first discuss the existing literature on word segmentation in infants, before zooming in on the current word segmentation study. Then I will elaborate on the method used to conduct the word segmentation experiment and present the results of the experiment. To conclude, the chapter will end with a discussion of the experimental results.

One of the first studies showing that infants use a rhythmic bias for word segmentation, which is consistent with the dominant stress pattern of their native language, was that by Morgan (1996), testing English-learning 6- and 9-month-olds. The study found that English-learning 9-month-olds preferred to segment disyllabic words with an SW stress pattern over disyllabic words with a WS stress pattern. This segmentation preference was not yet present at 6 months of age. According to the author, this implies that English-learning infants develop knowledge of what a rhythmically well-formed disyllabic word in their native language is between the ages of 6 and 9 months. The youngest age at which word segmentation based on language-specific rhythmic cues has been found is 7.5 months, also in English-learning infants

(Jusczyk, Houston & Newsome, 1999). In this study, the authors tested 7.5 and 10.5-month-old English-learning infants on their segmentation of SW and WS words from fluent speech. At 7.5 months of age, infants correctly segmented SW words and 'missegmented' WS words as SW words. These results point to an initial SW preference for segmenting words from fluent speech in English-learning infants based on the dominant word level stress pattern of English. Only three months later, at 10.5 months of age, they were also able to correctly segment WS words from fluent speech, a less common pattern in English.

Infants can also use phonotactic cues in word segmentation, this interacting with rhythmic cues. A study by Mattys, Jusczyk, Luce & Morgan (1999) investigated whether 9-month-old English-learning infants are sensitive to the phonotactic probabilities of consonant clusters appearing within or between words in the language they are learning. It appears that, indeed, English-learning infants of this age prefer to listen to SW non-words containing high-frequent within-word clusters over SW non-words containing high-frequent between-word clusters. However, when the stress pattern of the non-words was also manipulated, this influenced the preference of the infants. When infants listened to SW words, they again preferred non-words with high-frequent within-word clusters, thus chunking SW units, but when they listened to WS non-words they preferred words with high-frequent between-word clusters, thus splitting WS units. This suggests that stress leads the way: at 9 months of age infants rely more heavily on stress cues than on phonotactic cues for word segmentation. Yet, it also implies that infants of this age can already integrate multiple phonological cues, namely phonotactics and stress, in word segmentation. If they were unable to integrate these cues, we would not see the interaction between phonotactics and stress, but only the main effects of the cues separately. Hence, in a study investigating the use of stress cues on word segmentation, serious consideration should also be given to the phonotactic probabilities of the phoneme sequences in the stimuli used.

The results of Mattys et al. (1999) were confirmed in a segmentation study with 8-month-old English-learning infants conducted by Johnson & Jusczyk (2001). First they found that English-learning infants of 8 months of age can use only statistical cues, such as transitional probabilities, for word segmentation (which is a replication of the results found in an earlier study by Saffran et al., 1996). They then added an extra dimension to the experiment, namely stress, which conflicted with the statistical cues. In this case, 8-month-old English-learning infants weighed the stress cues more heavily than the statistical cues by segmenting the prosodically well-formed SW non-words instead of the statistically well-formed non-words. Thus, it seems as if English-learning infants of both 8 and 9 months of age heavily rely on stress cues for word segmentation.

Mattys et al.'s (1999) and Johnson & Jusczyk's (2001) results were in turn replicated by Thiessen & Saffran (2003), who also found that when stress cues and statistical cues collide English-learning infants at 8.5-9 months of age ignore statistical cues and use the SW stress pattern as a cue to word segmentation. However, the authors also tested 6.5-7-month-old English-learning infants with the same experiment and, contrary to the older age group, the younger infants did not ignore the statistical cues and use the stress cues, but rather seemed to ignore the stress cues and use the statistical cues only. All of these results seem to suggest that English-learning infants between 6 and 7 months of age do not use stress as a cue to word segmentation, whereas from 7.5 months onward English-learning infants do use the dominant stress pattern of their native language for word segmentation.

The results of the English-learning infants demonstrate the use of stress cues for word segmentation based on grouping principles: SW units are regarded as word-like units by English-learning infants, whereas WS units are not regarded as possible word-like units initially, but only later in development, around 10.5 months of age. However, the perspective I would like to adopt here is one in which a stressed syllable is regarded as a metaphorical 'anchor' in segmenting the continuous speech stream, a view similar to the approach taken by van Ommen (2016). A stressed syllable can, in this view, cue different positions of word boundaries depending on the dominant stress pattern of a particular language. For English, a strong syllable often marks a word onset, thereby alerting native listeners to the start of a new word and, hence, facilitates word segmentation. As German and Dutch are very similar to English in their word level rhythmic structure (SW), I expect infants learning these languages to employ a similar segmentation strategy as English-learning infants, and thus to employ a segmentation strategy in which a stressed syllable is regarded as a word onset, at similar ages.

According to Höhle (2002) for German and Houston, Jusczyk, Kuijpers, Coolen & Cutler (2000) for Dutch, infants learning these languages indeed segment SW words at 8 and 9 months, respectively. However, Kuijpers, Coolen, Houston & Cutler (1998) also report that Dutch-learning infants do not employ a segmentation strategy based on SW units yet at 7.5 months of age, while English-learning infants of this age do (Jusczyk et al., 1999). A possible explanation given by the authors for this discrepancy is the fact that the correspondence of full vowels to stressed vowels is less reliable in Dutch than it is in English. In English most unstressed syllables contain reduced vowels (see for exceptions the English rhythm rule: Hammond, 1984; Kaisse, 1987;

Gussenhoven, 1991), however, in Dutch an unstressed syllable can still contain a full vowel (Kager, 1989; Trommelen & Zonneveld, 1989). Furthermore, acoustic measurements of infant-directed speech (IDS) in American-English and Dutch also demonstrate that the difference between stressed and unstressed syllables is more exaggerated in American-English IDS than it is in Dutch IDS, mostly on the pitch dimension (Kuijpers et al., 1998). This may provide American-English infants, which all of the English-learning infants mentioned above are, with more salient stress cues than Dutch IDS provides for Dutch-learning infants.

All of these studies describe trochaic stress-timed languages. French is the only non-stress-timed language properly investigated for its prosodic cues in early word segmentation by infants. As French is a syllable-timed language, the unit of segmentation is not necessarily expected to be a disyllable rhythmic unit, but the syllable itself. Gout (2001) indeed found segmentation of monosyllabic words by 7.5 months of age; segmentation of disyllabic words is not found up to 11 months of age for infants learning European-French. However, for infants learning Canadian-French, Polka, Sundara & Blue (2002) did find 'language-specific' segmentation of disyllabic WS words by 7.5 months of age (see also Polka & Sundara, 2003; Polka & Sundara, 2012).

Given these mixed results for French-learning infants, Nazzi et al. (2006) ran a series of experiments to find out whether infants learning European-French can also segment disyllabic words at 8 months of age or whether they only segment one of the syllables from these disyllabic words, which would make more sense from an RCH perspective. They first tested French-learning 8-month-olds on their segmentation of disyllabic words and did not find evidence for segmentation at this age. They then tested 12-month-olds on their segmentation of monosyllables, in which they succeeded. At this age, French-learning infants can segment the individual syllables of disyllabic words, but not the words as a whole. Only at 16 months of age are French-learning infants able to segment disyllabic words from fluent speech. The authors view this as evidence in favor of the RCH. However, as we know from the discrimination and preference studies with French-learning infants discussed in Chapters 2 and 4, there is no phonological representation of stress at the word level in French-learning infants. Hence, properly investigating the influence of the position of word level stress on word segmentation requires a language which is also syllable-timed, but has final stress at the word level. According to the literature (Göksel & Kerslake, 2005), Turkish is such a language.

7.1.1 The current study

This chapter reports on an experiment in which I tested monolingual Dutch- and Turkish-learning 8-month-olds on their use of stress cues for word segmentation cross-linguistically. In this experiment, infants are tested on their segmentation of monosyllabic words (taking into account the syllable-timed nature of Turkish) from different contexts containing trisyllabic strings with different stress patterns preceding the target words. I chose to work with trisyllabic preceding contexts in order to create the three conditions presented in Table 7.1. The position of the stressed syllable in these trisyllabic strings was either initial (SWW), medial (WSW) or final (WWS) and the strings were always immediately followed by a monosyllabic target word, which always received stress. In actuality, in the design the SWW-S pattern takes the form of a control condition, as this pattern containing a so-called 'lapse' (two adjacent unstressed syllables) does not facilitate segmentation of the following monosyllabic target word for either language group. Yet, the WSW-S and WWS-S patterns were predicted to have different segmentation effects in the Dutch and Turkish groups, respectively.

The design of this experiment was set up as part of a larger cross-linguistic project on speech segmentation and the initial hypotheses were also linked to this design. Again, the SWW-S condition was introduced as a control condition, as I did not expect it to benefit either Dutch- or Turkish-learning infants in segmenting the target word. However, it was hypothesized that Dutch-learning infants would be aided in their segmentation of the target word in the WSW-S condition, as they would segment the SW sequence in the preceding context as a word-like unit and expect a new word to start at the next strong (S) syllable, which is the target word in this design. This particular facilitating effect of stress on segmentation would be interpreted as language-specific only if this condition would not facilitate segmentation of the target word for Turkish-learning infants. For Turkish-learning infants the expectation was that word segmentation would be facilitated in the WWS-S condition, as they would segment the WS sequence in the preceding context as a word-like unit and expect a new word to start immediately after. Again, whether this effect can be interpreted as language-specific is dependent on whether this particular condition effect is only found in the Turkish-learning group and not also in the Dutch-learning group.

For the WWS-S condition, it is actually plausible to find a facilitating effect on segmentation across languages, because this condition contains a so-called 'stress clash': two adjacent stressed syllables. Stress clash is assumed to be a highly marked phenomenon which is avoided in many of the world's languages by phonological processes such as stress retraction (Liberman & Prince, 1977). These phonological processes ensure that within a phonological word there is only one (primary) stress (or in the case of secondary stress, that the two stressed syllables are not adjacent). This assumption is supported by the use of stress clash as a universal constraint in Optimality Theory, which was formulated initially by Liberman (1975), Liberman & Prince (1977) and later adapted by Prince (1983), Selkirk (1984), Hammond (1984), Hayes (1984) and Hyde (2002) as a *Clash constraint which discourages adjacent stressed syllables. The importance of stress clash is further strengthened by the existence of the Unique Stress Constraint (USC) proposed by Gambell & Yang (2006). The USC states that a word can bear at most one primary stress (a strong syllable) and, consequently, that if two strong syllables are adjacent, a word boundary is postulated between them. The USC is argued to be an innate or universal constraint by the authors and it has proven its effectiveness in computational models of word segmentation by infants (Gambell & Yang, 2006).

Notice, however, that the results obtained in the preference experiments with Dutch- and Turkish-learning infants as described in Chapters 4 to 6 add hypotheses to the initial ones stated above and illustrated in Table 7.1 below. Nevertheless, the design of this word segmentation experiment remains suitable for testing these hypotheses. Crucially, the results from the preference experiments (in Chapters 4 and 5) show a lack of evidence for a language-specific word level WS preference in the older Turkish-learning infants. The only real preference found in the Turkish group was a preference for the SW pattern at 4 months of age, but this preference was interpreted as reflecting an early preference for the phrase level stress pattern of Turkish. Hence, I either expect Turkish-learning infants not to use stress cues for word segmentation at all, similar to French-learning infants, or I expect them to use more 'universal' cues for segmentation, such as stress clash. This distinction can, of course, only be made against the background of the findings for the Dutch-learning infants in this experiment: if they do not demonstrate the use of stress clash for segmentation, then it is language-specific for the Turkish-learning infants. However, if the Dutch-learning infants do use stress clash as a segmentation cue, I interpret it to be a universal cue. Taking the preference results into account in a modification of my initial expectations, I still predict that the WSW-S condition facilitates segmentation for the Dutch group, but not for the Turkish group. Moreover, I do not predict a language-specific facilitation of the WWS-S condition for the Turkish-learning infants anymore. It is, however, plausible that I find a cross-linguistic facilitation of segmentation in this last condition for both language groups due to the stress clash embedded in this condition. The expectation regarding the lack of a facilitating effect of the SWW-S condition for both language groups remains unchanged. These predictions are represented in Table 7.1.

Table 7.1: *The hypotheses for the word segmentation experiment per language group and per condition (LS = language-specific, UN= universal).*

Condition	*Dutch*	*Turkish*
SWW-S	No segmentation	No segmentation
WSW-S	LS segmentation	No segmentation
WWS-S	(UN segmentation)	(UN segmentation)

7.2 Method

7.2.1 Participants

The participants in this experiment were 24 Dutch-learning and 24 Turkish-learning 8-month-olds, who (for both language groups) were the same 8-month-old infants first tested with the preference experiment described in Chapter 5 and then with the segmentation experiment (always in this order). The age of the participants was crucially determined by the finding in the literature that 7.5 months is the youngest age at which word segmentation abilities based on stress cues have been found (Jusczyk et al., 1999). For this reason, I only tested infants on their word segmentation abilities in the 8-month-old age groups and not younger. The mean age of the Dutch-learning 8-month-olds is 252 days and the mean age of the Turkish-learning infants is 256 days. All participants were healthy full-term infants without any known hearing or visual impairments and were being raised in a monolingual Dutch or Turkish environment. An additional 3 and 6 infants were tested in the Dutch and Turkish language groups, respectively. Data from these infants could not be included due to crying, fussiness or technical issues. Participant information is summarized in Table 7.2.

Table 7.2: *Included and total number of infants tested per language group with mean age and age range in days.*

Language group	*Dutch*	*Turkish*
Included (total)	24 (27)	24 (30)
Mean age (range)	252 (237-274)	256 (236-282)

7.2.2 Material

Differently from the previously described preference experiments, which measured spontaneous preferences without any specific exposure in the lab, this experiment is an Artificial Language Learning (ALL) experiment containing a familiarization phase. In an ALL experiment it is, hence, possible to test word segmentation cross-linguistically by using the same linguistic material for different language groups. The experiment thus consisted of two parts: a familiarization phase and a test phase. In the familiarization phase infants heard three passages consisting of six phrases (divided by a 250 ms pause) per passage. Each phrase consisted of a variable six-syllable string built from CV(C) syllables which did not contain any embedded words, except for some low-frequency monosyllabic words which were assumed to be unknown to the infants, in either Dutch or Turkish. Each six-syllable string did contain a monosyllabic (CVC) target pseudo-word which was the same word for each phrase within a passage. The syllables of the strings were created using Python and from a limited set of phonemes (only phonemes which are part of the phoneme inventories of both languages) with constraints on transitional probabilities (in order to avoid phonotactic cues to segmentation, such as lower TPs between the last syllable of the context and the target words than between the context syllables itself) and vowel harmony (for Turkish). The segmental structure of the material in the familiarization phase is represented in Table 7.3.

Each passage had its own particular stress pattern on a three-syllable string preceding the monosyllabic target word in each phrase: either strong-weak-weak (SWW), weak-strong-weak (WSW) or weak-weak-strong (WWS). Two filler-syllables were used on the edges of the six-syllable strings to vary the position of the target word within the phrases by varying the position of one or both filler-syllables between the onset and offset of the strings. Two different versions of the experiment were created, each with different target words embedded in different phrases. This was done, because half of the infants already heard the word /dɑr/, which is part of version 2, in the stimulus /dɑrnɑm/ presented in the preference experiment (Chapter 5). Furthermore, within each language group within each version there were four different lists for which each target word was connected to a different preceding stress pattern in order to avoid segmentation preferences based on the properties of the target words themselves. For each list a different target word, and with that, one of the four passages, did not appear in the familiarization phase: it was therefore by definition a novel word in the test phase. A familiarization phase, for example, consisted of the first three passages of version 1 (the /bɛr/-,

/dim/-, and /sum/-passage), whereas in the test phase all four target words were presented as word lists (/bɛr/, /dim/, /sum/ and /fɑr/). The prosodic structure of the material in the familiarization phase is represented in Table 7.4.

Table 7.3: *The material (in IPA[1]) of the two versions of the segmentation experiment's familiarization phase. The underlined parts of the stimuli represent the preceding context, the non-underlined parts the filler syllables and the bold parts the target words.*

Version 1	*Version 2*
bi se mi **bɛr** ne de	mi fe fi **bɛm** de bi
re me le **bɛr** si fi	re bi de **bɛm** si le
se ri fi ne **bɛr** bi	si re de ri **bɛm** bi
fe li fi ri **bɛr** de	bi me ne li **bɛm** le
di si mi li re **bɛr**	ri me si fi di **bɛm**
fi le me ri di **bɛr**	ne di mi re le **bɛm**
si de fe **dim** bi re	fu ba na **dɑr** sa fa
mi le ri **dim** li fi	ra lu du **dɑr** nu mu
fe si mi ni **dim** se	ru ba fu bu **dɑr** ma
fe le be ne **dim** de	nu ma fa la **dɑr** ba
me ri ni li be **dim**	sa nu su la lu **dɑr**
re ni de mi me **dim**	ba ru mu nu ra **dɑr**
du ra ru **sum** bu sa	fu ma da **sur** du ru
ra ma da **sum** lu ru	fa ma fu **sur** bu nu
bu nu na fa **sum** sa	la nu lu bu **sur** sa
ma ru lu sa **sum** du	mu du fu fa **sur** na
ra bu fa da fu **sum**	da bu mu ru sa **sur**
lu mu bu ba ra **sum**	fu nu sa ru mu **sur**
bu ru da **fɑr** na ma	mi ne re **fim** be ni
nu fu ba **fɑr** su ru	ri ne mi **fim** li ni
ra mu la nu **fɑr** ba	si ne fe li **fim** de
ru du lu na **fɑr** mu	ri be mi di **fim** si

[1]The exact realization of the phonetic and phonological variation within and between the two languages is not reflected in this transcription. The /r/, for example, is realized differently in initial and final positions in Dutch, but is also different from the /r/ in Turkish.

fu bu ba su du **far**	bi be de me le **fim**
bu nu lu da la **far**	se fe mi de ni **fim**

Table 7.4*: The prosodic structure of the phrases in the familiarization passages (context and target word in capitals, target word additionally in bold).*

Passages	*1 - SWW*	*2 - WSW*	*3 - WWS*
Phrase 1	SWW**S**ww	WSW**S**ww	WWS**S**ww
Phrase 2	wSWW**S**w	wWSW**S**w	wWWS**S**w
Phrase 3	wwSWW**S**	wwWSW**S**	wwWWS**S**
Phrase 4	SWW**S**ww	WSW**S**ww	WWS**S**ww
Phrase 5	wSWW**S**w	wWSW**S**w	wWWS**S**w
Phrase 6	wwSWW**S**	wwWSW**S**	wwWWS**S**

The stimuli were created from natural speech spoken by a highly competent multilingual speaker of Dutch and Turkish, who recorded the Dutch and Turkish version of the stimuli on different days. The six-syllable strings were recorded as whole phrases in which the target-word always had natural stress as realized by the speaker, whereas the different stress patterns on the trisyllabic context preceding the target word were created by resynthesis using a script in Praat (Boersma & Weenink, 2013). This would ensure that each string was segmentally identical in each condition, while differing in their stress pattern. With the script, the preceding context was first made prosodically flat and also isochronous by giving the rhyme of each syllable the mean duration of the three syllables in the context. After that, the stressed syllable's duration was multiplied by a ratio of 1.5, the pitch was increased by 8 semitones (the pitch-peak was in the first quarter of the rhyme) and the intensity was boosted with 8 dB (van Ommen, 2016).

Due to this resynthesis, I am certain that the stressed syllable in the preceding context was always longer in duration and higher in pitch and intensity than the unstressed syllables in the context. However, acoustic measures were taken (also using Praat) to check for whether the two stressed syllables in the string (1: context, 2: target word) were equal or different in their duration, pitch and intensity. The duration differed only in the Turkish version of the stimuli ($t(1,286) = -2.019$, $p=.045$), with a longer duration of the second stressed syllable (naturally produced on the target word), but not in the Dutch version of the stimuli. The pitch differed in both language versions of the stimuli ($t(1,286) = -16.404$, $p=.000$ and $t(1,286) = -6.356$, $p=.000$, for Dutch and Turkish respectively), with the higher pitch always on the second stressed syllable. The intensity also differed in both language versions (t(1,286) = 10.883, p=.000 and t(1,286) = 9.653, p=.000, for Dutch and Turkish respectively), yet with the higher intensity always on the first stressed syllable, probably due to the resynthesis which consisted of an 8 dB intensity boost on the this syllable in the context. A summary of the acoustic measures can be found in Tables 7.5 and 7.6.

Table 7.5: *The average values of the Dutch version for duration, mean pitch and mean intensity of the rhyme part of the first and second stressed syllable with their standard deviation (SD).*

Syllable type	*Duration (ms)*	*Pitch (mHz)*	*Intensity (mdB)*
Stress 1 (SD)	123.52 (23.64)	341.68 (32.24)	84.67 (1.79)
Stress 2 (SD)	128.22 (27.06)	423.24 (50.20)	81.18 (3.41)

Table 7.6: *The average values of the Turkish version for duration, mean pitch and mean intensity of the rhyme part of the first and second stressed syllable with their standard deviation (SD).*

Syllable type	*Duration (ms)*	*Pitch (mHz)*	*Intensity (mdB)*
Stress 1 (SD)	112.87 (19.69)	365.33 (38.14)	83.93 (2.15)
Stress 2 (SD)	119.60 (34.80)	402.72 (59.41)	80.94 (3.03)

In the test phase the infants listened to four word lists: each word list consisted of multiple tokens of the same monosyllabic word with a 750 ms pause in between the words. Three lists contained monosyllabic words which occurred in the familiarization phase and could potentially have been segmented by the infants, whereas one word list contained repetitions of a novel word, which did not occur in the familiarization phase. As stated above, this novel word also differed per infant as the infants were divided over the four lists within the two versions. The word lists were created by cutting the target words from the phrases of the familiarization phase. In other words, the isolated words in the test phase had exact acoustic matches with the embedded target words in the familiarization phase.

7.2.3 Procedure

Similar to both preference experiments described in Chapters 4 and 5, the Central fixation auditory preference procedure (Cooper & Aslin, 1990) using eye tracking was employed to conduct this study. The same hardware and software and test situations as in the preference experiment described in Chapter 5 held for the word segmentation experiment reported here. The word segmentation experiment started with an attention grabber which was immediately followed by the familiarization phase. During the familiarization phase the visual stimuli consisted of three colorful pictures (drawings) of objects or animals, appearing in different positions on the screen. The pictures and their position changed with every new phrase in the auditory stimuli. The infants heard each of the three passages containing six phrases twice in a pseudo randomized order (the same passage was never immediately repeated) and the infants thus had twelve segmentation opportunities per target word during the familiarization phase. The familiarization phase had a fixed duration and was not infant controlled, but their attention during familiarization, in terms of looking time towards the AoI (in this experiment the part of the screen in which the three images appeared) was measured. Hence, for all infants the duration of the familiarization phase was the same, but the total duration of their attention towards the stimuli in the familiarization phase could be different.

After the familiarization phase, the attention grabber appeared again and was followed by the first test trial. During the test phase, the visual stimulus was again a picture of a smiling female face; however, the picture was different from the one in the preference experiments (Figure 7.1). The test phase consisted of 8 trials: each of the four word lists were played twice in a pseudo-randomized order (the same word list was never immediately repeated). The duration of each test trial and of the test phase as a whole was dependent on the interest of the infant. In other words, the test phase was infant-controlled, which means the test trials ended when the infant looked away from the visual stimulus for more than 2 consecutive seconds.

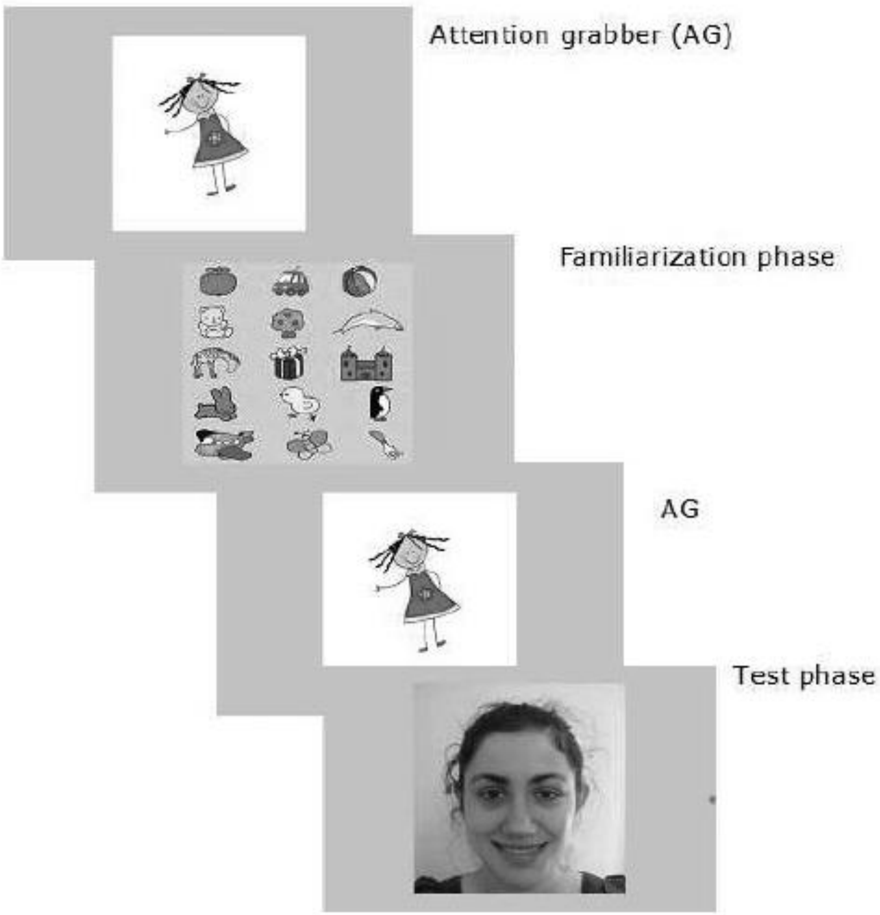


Figure 7.1: *The outline of the word segmentation experiment using eye tracking.*

7.3 Results

Similar to the analysis of the preference experiment described in Chapter 5, several scripts were run on the raw data resulting from the ZEP output in order to extract the information needed to calculate the Total Looking Time (TLT) per participant per test trial. The temporal and spatial criteria for 'looking' and 'looking away' were also the same as in the second preference experiment. The participant inclusion criteria for the word segmentation experiment stipulated that only infants who completed the whole experiment, both the familiarization phase and the test phase, were to be included in the analysis. For all infants from both language groups combined, the mean TLT during the novel word trials (new) is 7512 ms (SD = 4926 ms), the mean TLT during the SWW (condition 1) word trials is 6575 ms (SD = 4924 ms), during the WSW (condition 2) word trials the mean TLT is 8088 ms (SD = 5632 ms) and during the WWS (condition 3) word trials 6430 ms (SD = 4879 ms). These mean TLTs per condition are presented in Figure 7.2 (see also Figures 7.3 and 7.4). The mean TLTs per language group and the average for both language groups together are shown in Table 7.7.

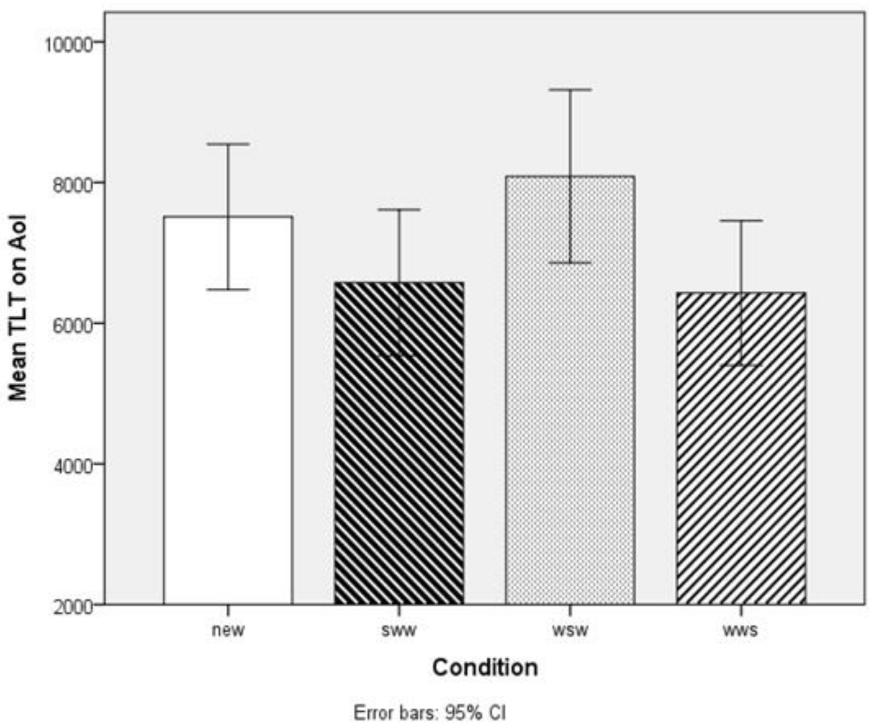


Figure 7.2: *The mean TLT in ms in the test phase for all infants from both language groups and per condition.*

Table 7.7: *The mean TLT (SD) in ms in the test phase per language group and per condition.*

Conditions	*0 (new)*	*1 (SWW)*	*2 (WSW)*	*3 (WWS)*
Dutch	7151 (5051)	6315 (4898)	7418 (5714)	5439 (4940)
Turkish	7857 (4834)	6818 (4990)	8743 (5540)	7316 (5525)
Average	7512 (4926)	6575 (4924)	8088 (5632)	6430 (4879)

The purpose of the current experiment is to see whether Dutch-learning and Turkish-learning infants use (language-specific) stress cues for word segmentation at 8 months of age. In order to determine whether there were significant effects of the variables condition and language group, I ran a statistical analysis. As the data were not normally distributed, I first ran a normal transformation on the data using Blom's formula (Blom, 1958) and then ran the analysis on the normalized TLTs. I ran a Linear Mixed Model with language and condition as fixed factors, participant and trial as a random factors and mean TLT in the familiarization phase as a covariate, as I found a significant positive correlation between the variables TLT on the AoI during familiarization and TLT on the AoI during the test phase ($r = .186$, $p=.000$).

I indeed found a significant main effect of the factor TLT in familiarization ($F(1,47) = 5.902$, $p=.019$) and, more importantly, a significant main effect of the factor condition ($F(3,307) = 2.660$, $p=.048$). However, there is neither a significant main effect of the variable language group ($F(1,47) = 0.881$, $p=.353$), nor a significant interaction of the variables language group and condition ($F(5,307) = 0.800$, $p=.629$). This means the language groups behaved similarly overall in the experiment and also with respect to the variable condition. Therefore, I cannot confirm the use of language-specific cues for word segmentation at this age. The mean TLTs per condition per language group are presented in Figures 7.3 and 7.4 for the Dutch and Turkish language group respectively.

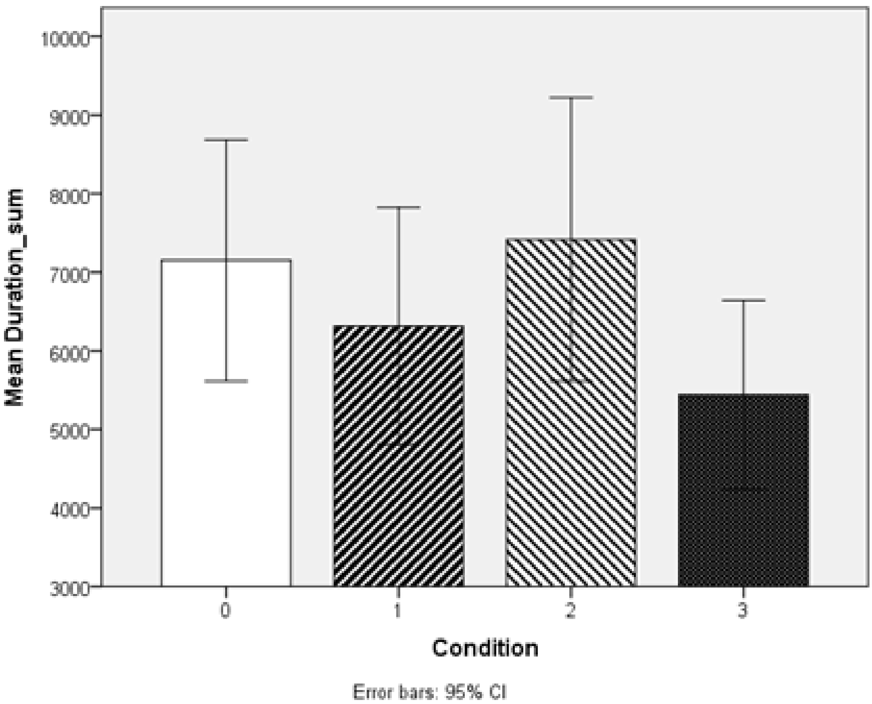


Figure 7.3: *The mean TLT in ms in the test phase for the Dutch-learning infants per condition (0 = new target word, 1 = SWW, 2 = WSW, 3 = WWS).*

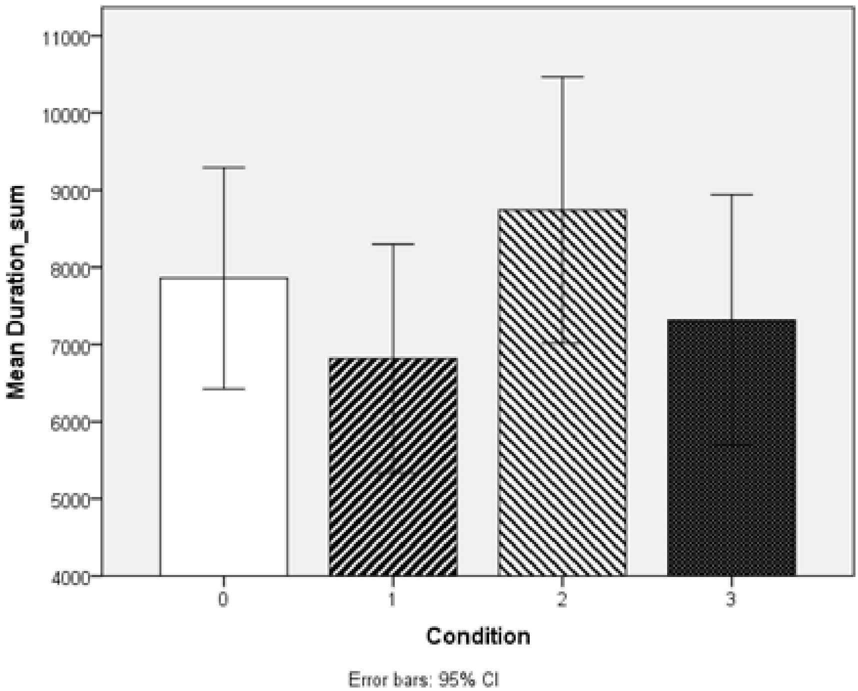


Figure 7.4: *The mean TLT in ms in the test phase for the Turkish-learning infants per condition (0 = new target word, 1 = SWW, 2 = WSW, 3 = WWS).*

As I found a main effect for the factor condition, we can look at the pairwise comparisons between the different conditions. I found a significant difference between the words from the WSW condition and the words from both the SWW and WWS conditions (p=.028 and p=.014, respectively). However, I did not find a significant difference between the words from the WSW condition and the novel words. This means that the infants looked longer during the words from the WSW condition than during the words from the SWW and WWS conditions, but looked just as long during the words from the WSW condition as during the novel words. In other words, they preferred novel words and words from the WSW condition over the ones from the SWW and WWS conditions. I interpret this preference as a novelty preference, meaning they actually segmented the words from the SWW and WWS conditions and not the words from the WSW condition. As they recognized the words from the SWW and WWS conditions, they were less interested in those words than they were in the novel words and in the – to them – novel words from the WSW condition, which they seem not to have segmented during the familiarization phase. This, again, is a surprising result, which I will further discuss in the section below.

7.4 Discussion and conclusion

Based on earlier segmentation experiments conducted with English-, German- and Dutch-learning infants around 8 months of age (Morgan, 1996; Jusczyk et al., 1999; Mattys et al., 1999; Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003; Höhle, 2002; Kuijpers et al, 1998; Houston et al., 2000), I expected Dutch-learning infants of the same age to also use their native trochaic word stress pattern for word segmentation. However, the results found here actually point in the opposite direction: the trochaic WSW pattern does not facilitate segmentation of the following target word, because the word embedded in this condition in familiarization was just as novel to the Dutch-learning infants as the actual new words. In fact, it seems WSW is the only condition in which the infants do not segment the following target word, whereas they do segment the target words in the other two stress conditions, namely SWW and WWS. For the Turkish-learning 8-month-olds, I initially hypothesized that the WWS stress pattern would be (the most) facilitating for word segmentation, because a stressed syllable signals a word-offset in Turkish and based on previous mixed results from several studies testing French-learning infants around this age. Some of these studies only found segmentation of (strong) monosyllabic words (Gout, 2001; Nazzi et al., 2006), while others found segmentation of iambic WS words (Polka et al., 2002; Polka & Sundara, 2003; Polka & Sundara, 2012) by the age of 8 months.

However, given the findings which did not confirm a word level WS representation for the older Turkish-learning infants in the preference experiments described in Chapters 4 and 5, I did not expect to find language-specific stress effects in a word segmentation experiment for Turkish-learning infants. Consequently, if anything, I also expected Turkish-learning infants to benefit from the WSW pattern in segmentation, although they could of course still segment the target word in the WWS condition based on the stress clash in this condition cueing a word boundary. If Dutch-learning infants would also segment the target word in this last condition, this would be an additional reason to discard a language-specific interpretation of a potential facilitating effect of the WWS pattern for word segmentation by Turkish-learning infants at 8 months of age.

I indeed did not find language-specific effects in the word segmentation experiment: Dutch- and Turkish- learning 8-month-olds demonstrate similar behavior in this type of word segmentation task. Hypothetically, an explanation for this might be that the task required the segmentation of rather simple, monosyllabic word forms and they could do this independently of the stress pattern preceding these words (see also Juszcyk & Aslin, 1995). However, if this was the case, I would not find differences between stress conditions, but only differences between the novel words and all three familiarized words. As I did find differences between words from different stress conditions, there is definitely an effect of stress pattern on segmentation. Nevertheless, the direction of this effect is unexpected: I expected to find a facilitating effect for words following the trochaic WSW pattern. As the looking times during the words from this condition were similar to the looking times during the novel words, I can only interpret this as suggesting that these words were not actually segmented during the familiarization phase.

By the same reasoning, the significantly shorter looking times during the words from the SWW and WWS conditions indicate that these words were in fact segmented during the familiarization phase. For the latter, the WWS condition, a language-independent facilitating effect can be explained by the notion of stress clash. As this condition contained two adjacently stressed syllables in familiarization, it contained a stress clash and this is arguably a universal cue to word boundaries. Independently of the language the infants are learning, at 8-months, they use this cue to place a word boundary between the last syllable of the context and the target word, and with that, they segment the monosyllabic target word embedded in this stress condition.

Language-independent segmentation of the monosyllabic target word following the SWW stress pattern is more complex to account for theoretically. We cannot really rely on the literature to date for this, as most of the literature reports on the segmentation of disyllabic target words. In these studies there is no distinction between a trochaic word ((W)SW) and a non-trochaic, but also initially stressed word (SW(W)). In order to explain the segmentation of monosyllabic target words in this last condition we actually have to delve into the literature describing the segmentation of trisyllabic words. I will first look at several studies investigating adult word segmentation and then also at data from infants segmenting trisyllabic words.

In a study investigating cross-linguistic differences in the use of different cues for speech segmentation Tyler & Cutler (2009) tested English, Dutch and French adult native speakers on their use of different cues, namely (i) only transitional probabilities, (ii) only durational cues or (iii) only pitch cues at both the left and right edges, to segment the speech stream in an ALL paradigm. The participants were supposed to segment either three or four syllable words in each of the three conditions. The results show a facilitating effect of durational cues on segmentation at the right edge (word final) only and for all three language groups. The authors hence concluded that this is a universal cue to word endings. In the pitch condition, the English participants only used the left-edge (word initial) cue, the French participants only used the right-edge (word final) cue, and the Dutch participants, interestingly, used the pitch cue at both edges. This result is explained by the authors by the fact that native Dutch speakers would rely more on pitch cues (such as spectral tilt) for signaling word boundaries than English speakers, who rely more on vowel quality, and therefore Dutch speakers can also use this cue to segmentation at the (also for Dutch) non-typical right-edge position. The results from the Dutch-learning 8-month-olds discussed above seem to confirm this idea of pitch signaling both word-onset and word-offset, as pitch cues are present in all stress conditions in my ALL word segmentation experiment and Dutch-learning infants treat them as the edges of trisyllabic 'words' as well.

On the other hand, Endress & Hauser (2010) claim that speech segmentation can take place without any language-specific knowledge of prosodic cues. They, for example, suggest that pitch is often used as a word initial cue, while duration is mostly used as a word final cue. This suggestion is, of course, related to the Iambic-Trochaic Law (ITL), which states that pitch movement marks unit-onsets, whereas lengthening marks unit-offsets as previously discussed in Chapter 6. However, as both of these cues are present in the stressed syllables of the word segmentation experiment reported on here, this alone cannot explain why the infants in this experiment would treat one stressed syllable as a word-onset and another stressed syllable as a word-offset. Moreover, for language, the universal mechanism of the ITL holds at the phrase level for most languages, but not always at the word level, as language-specific phonetic cues play a larger role at this level, such as Turkish only using pitch to mark the offset of a word, but not duration. Endress & Hauser (2010) also admit that is it more likely that their participants detected phonological phrase boundaries, rather than word boundaries.

Nevertheless, in convergence with positional information about the phrase edge (Endress & Mehler, 2009b) and, consequently, with the integration of prosodic and positional information, this combination of prosodic and positional information can possibly explain the universal use of stress cues for word segmentation as found in my experiment. In the first part of their study Endress & Hauser (2010) found that English-speaking adults could only segment words in an artificial language with French intonation when the target words where aligned with the, in this case, final edge of the intonation phrase and not when the target words were positioned in the middle of the intonation phrase and thus not aligned with a prosodic boundary cue. In my experiment the phrases in the familiarization passages had the prosodic structures displayed in Table 7.4.

Based on these structures, it is plausible that the infants in my experiment segmented the words from the SWW condition, because only in this condition (and never in the other two conditions) is a stressed syllable (underlined in Table 7.4) in 30% of the stimuli aligned with the start of a phonological phrase, a boundary which is clearly marked by the 250 ms pause between all phrases. As all phrase boundaries are (by definition) also word boundaries, this structure tells the infants that words in the artificial language presented to them in the familiarization phase (can) start with a stressed syllable. (In contrast, in the other two conditions, phrases always start with an unstressed syllable.) This information may have helped them in segmenting the next stressed syllable from the same phrase, which happened to be the target word in this experiment, and with that, facilitated word segmentation in this condition. In the WSW condition, this alignment with the phrase boundary was never present, as all phrases in this condition started with an unstressed syllable. For this reason, infants could not make use of the combination of prosodic and positional cues in this last condition.

Additional evidence for this explanation comes from Toro-Soto, Rodríguez-Fornells & Sebastián-Gallés (2007), who conducted a study with adult native speakers of Spanish, a language which also has a dominant SW stress pattern. The participants were tested on their word segmentation in an artificial language consisting of trisyllabic words based on statistical information in five different conditions: a flat condition (no stress), a stress initial condition (SWW), a stress medial condition (WSW), a stress final condition (WWS) and a random condition (no stress, no statistical information) as a control condition. The results show that Spanish speaking adults, surprisingly, do not use their native medial stress pattern (WSW) for segmentation. Contrary to this, the participants were hindered in their segmentation of the statistical trisyllabic words in the medial stress condition, whereas they performed well in the flat, stress initial and stress final conditions. These results are very similar to the results found in my word segmentation experiment with Dutch- and Turkish-learning infants for whom I hypothesized a facilitating effect of the stress medial (WSW) pattern as well. The authors also explain their results by the so-called 'edge effect': the idea that stress facilitates segmentation when it coincides with a linguistic boundary, because in this case the stress initial and final conditions perceptually enhanced the statistical word boundaries. The authors conclude that perceptual salience can under certain conditions overrule the use of language-specific stress patterns in a word segmentation task.

The studies described above all refer to experiments conducted with adult participants, but we can also find some evidence for segmentation by 9- and 7-month-old English-learning infants taking place based on an SWW pattern rather than a WSW pattern in a study of Curtin, Mintz & Christiansen (2005). Their experiment is very similar to the experiment by Toro-Soto and colleagues (2007), except that there are no statistical segmentation present in the speech stream. This is comparable to the random condition in Toro-Soto et al. (2007), but in the study of Curtin and colleagues, stress was placed on every third syllable in the speech stream. In both the 9- and 7-month-old age groups, the English-learning infants preferred to segment SWW words over WSW words and they made no distinction between WSW and WWS words. This result demonstrates a clear initial stress strategy for segmentation, but by no means a trochaic strategy. Although this result is more expected for English, having default antepenultimate stress, than for Dutch, it still shows the dominance of an initial stress strategy for segmentation over an, also plausible, trochaic segmentation strategy.

Moreover, the authors found a novelty preference for the novel (control) words and unsegmented stress medial (WSW) words, demonstrated by longer looking times for these words than for the segmented stress initial (SWW) words in the test phase. This result is very similar to the preference for novel words and unsegmented words from the WSW condition in my predictive word segmentation experiment and strengthens this interpretation of my data. Furthermore, Curtin and colleagues (2005) did not find any difference between control words, stress medial words (WSW) and stress final words (WWS), which shows that in their experiment infants did not segment WWS strings as words. This result, in turn, strengthens my interpretation of segmentation in the WWS condition in my own word segmentation experiment as segmentation due to the presence of stress clash, because in the Curtin et al.'s (2005) study there is no segmentation of this pattern in the absence of stress clash.

To conclude, this predictive word segmentation experiment testing Dutch- and Turkish-learning 8-month-old infants, does not demonstrate a facilitating effect of a WSW stress pattern when this pattern precedes a stressed target word for either language group. In fact, both language groups use an initial stress strategy for segmentation (provided the stressed syllable occasionally coincides with the beginning of a phonological phrase) and stress clash (an arguably universal cue to word boundaries) to predict an upcoming word boundary and, with that, segment the monosyllabic words following either one of these two stress patterns. Hence, I do not find any evidence for the use of language-specific stress cues in this type of word segmentation experiment by 8-month-old infants learning either Turkish or Dutch, but I do find evidence in the results from the current experiment for the use of 'edge' stress cues related to salience in segmenting an artificial language by infants of this age learning languages of different types.

This finding is, however, diametrically opposite to the results found for adult speakers of these same languages. Van Ommen (2016) demonstrated that adult Dutch speakers' word segmentation in a non-word spotting task in which participants had to spot a disyllabic pseudo-word also preceded by one of the following three stress patterns: SWW, SWS and WWS, was facilitated by their native (W)SW word stress pattern and adult Turkish speakers' word segmentation was facilitated by their native (W)WS word stress pattern. Therefore, infants must shift their segmentation strategy from using the arguably 'universal' cues discussed above to using their language-specific stress patterns to aid word segmentation at a later point in their language development. Future studies examining monolingual Dutch- and Turkish-learning infants' stress-based word segmentation strategies at older ages (from 9 months of age) can shed more light on when in development this shift towards adult-like language processing takes place.

Chapter 8
General discussion and conclusion

The overall aim of this study was to test the language-specificity hypothesis of metrical segmentation for native language acquisition in infants aged 4, 6 and 8 months in a cross-linguistic design. This was done by contrasting a (statistically predominant) initial stress language, Dutch, in which a weak-strong (WS) syllable sequence implies a discontinuity and a strong-weak (SW) sequence implies both continuity (must be parsed together) and discontinuity (a word onset precedes S), with as the mirror image, a (statistically predominant) final stress language, Turkish, in which SW implies discontinuity and WS implies both continuity (must be parsed together) and discontinuity (a word offset follows S). I hypothesized that infants analyze the unsegmented input statistically in terms of rhythmic bigrams: pairs of syllables (strong-weak or weak-strong) and the alignment of these pairs with phrase edges (initial or final) which are marked by universal characteristics.

Before investigating word segmentation, it first had to be established that infants learning these two languages demonstrated having a native rhythmic representation of the dominant stress pattern of their language from 6 months of age. This was done by testing rhythmic preferences in the two experiments reported on in Chapters 4 and 5. The first general hypothesis for these experiments was that infants' language-specific preferences for rhythmic units develop around 6 months of age as a function of the distributional patterns in the input they receive. The idea was thus that infants first represent this distributional knowledge as under- or overrepresented bigrams, before extracting a rhythmic unit (SW or WS) when one of the bigrams is overrepresented. This approach offered a new account of the development of language-specific rhythmic preferences in infants stating that the acquisition of a rhythmic unit is based on continuous speech, instead of a protolexicon (e.g. Swingley, 2005), and it thereby complemented previous accounts related to rhythm class, as proposed by Nazzi & Ramus (2003).

To test this general hypothesis, I investigated Dutch-learning infants who were expected to show an SW preference at 6 months of age, similar to German-learning infants (Herold et al., 2008), and Turkish-learning infants who were expected to show a WS preference at 6 months of age. In Dutch, SW bigrams are overrepresented (at the left edge) compared to WS bigrams, whereas in Turkish WS bigrams are overrepresented (at the right edge) compared to SW bigrams (van Ommen, 2016), hence, demonstrating a reverse distribution. The research question for the first experiment was: do Dutch-learning infants and Turkish-learning infants show a language-specific rhythmic preference, and if so, at what age?

I predicted that Dutch-learning and Turkish-learning 4-month-olds would fail to show a bias, whereas at 6 months of age both Dutch and Turkish language groups would show a native SW and WS preference, respectively. The results for the Dutch-learning infants were as expected: I indeed did not find a strong rhythmic preference in the 4-month-olds and I did find a clear SW preference in the 6-month-olds. The results for the Turkish-learning infants were, however, opposite to the expectations: I found a clear SW preference in the 4-month-olds, but no strong preference in the 6-month-olds. I attempted to account for these data with the Universal Trochaic Bias (UTB) hypothesis, but this hypothesis cannot explain the full picture also emerging from the literature: initial trochaic biases are rarely found in perception studies testing infants younger than 6 months.

An alternative hypothesis was proposed at the end of Chapter 4, which led to the second preference experiment, a follow-up experiment, which was described in Chapter 5. This alternative hypothesis was inspired by a study on Hebrew-learning 9-month-olds, who demonstrate a native WS preference when listening to a Hebrew speaker, but a non-native SW preference when listening to an English speaker. As the infants in first preference experiment were also not listening to a speaker of their native language, but to a Spanish-speaker, their overall SW preference was hypothesized to be caused by not perceiving their native phonetic cues to stress, as formulated in the Native-language dependency (NLD) hypothesis. However, with the follow-up experiment testing this hypothesis, I replicated the results from the first preference experiment, and with that, concluded that the initial results could not be due to the speaker delivering the stimuli.

This conclusion led to the search for a more parsimonious account explaining the data from my own experiments as well as the data from the literature. This quest was described in the interim discussion in Chapter 6 and ended in an account combining the Iambic-Trochaic Law and the phrasal prominence proposal. The main assumption of this combined account is that infants before 6 months of age process all spoken stimuli as phrases as they are considered to be prelexical at this stage in development and therefore do not the distinguish between phrases and words yet. When infants become lexical (after 6 months of age) they start distinguishing words from phrases and also develop different representations for these two levels of linguistic processing.

More specifically, representations of prominence relations can differ between the word and the phrase level as languages can have different dominant rhythmic patterns for each of these levels of which the latter is ultimately related to word order. Dutch has both SVO and SOV word orders for main and subordinate clauses, respectively. These word orders may also be differentially marked by phonetic cues, as was found for German (Nespor et al., 2008), with final duration marking SVO clauses creating a WS rhythmic pattern at the phrase level, and initial pitch marking SOV clauses creating an SW rhythmic pattern at the phrase level. Thus, for 4-month-old Dutch-learning infants who are still phrase level 'processors', SW and WS patterns are equally frequent and equally well-formed. Therefore, the finding that they do not demonstrate a strong rhythmic preference at this age in the first experiment becomes much less surprising.

In Turkish, however, the most frequent word order is SOV and thus at the phrase level, an SW rhythmic pattern is dominant. This can explain the clear SW preference in Turkish-learning 4-month-olds. When infants start building up a lexicon from 6 months of age (Bergelson & Swingley, 2012), they start to differentiate between phrases and words, and discover that word level prominence patterns can differ from phrase level patterns. In the case of Dutch, the frequency of processing SW patterns increases as this is the most common stress pattern of Dutch words, leading to an overall dominant SW pattern. This is reflected in a clear SW preference found in the Dutch-learning infants at 6 months of age in both preference experiments reported on in Chapters 4 and 5. For Turkish, differentiating between the phrase and the word means that the dominant SW phrase level pattern gets competition from a dominant WS word level pattern, thereby cancelling out a single dominant pattern overall. This is, in turn, reflected in the lack of a strong rhythmic preference in Turkish-learning infants at 6 months of age.

The second general hypothesis was that this distributionally-based knowledge of rhythmic patterns also underlies infants' metrical segmentation abilities which were expected to develop around 8 months of age. Once a certain rhythmic unit is extracted around 6 months of age, it can subsequently be used to inform word segmentation. That is, a transition within an overrepresented bigram is interpreted as a continuity cue, while a transition within an underrepresented bigram is interpreted as a discontinuity or segmentation cue. If infants of both language groups would have demonstrated these language-specific preferences at 6 months of age, I would have expected to find the use of these language-specific stress patterns for word segmentation at 8 months of age. Thus, Dutch-learning infants would regard a transition from a strong to a weak syllable (SW, 'trochaic') as a continuity cue, which leads to chunking of SW-units, while a transition from a weak to a strong syllable (WS, 'iambic') would be seen as a discontinuity cue, which leads to splitting of WS-units in a word segmentation task. For Turkish-learning infants the opposite pattern should have held: SW-units are split based on the thereby provided discontinuity cue, while WS-units are chunked, based on the continuity cue they provide in Turkish.

However, as I did not find language-specific preferences for both language groups at 6 months of age, this motivated further research steps. In these, the second general hypothesis and the research questions related to it were slightly altered. Consequently, the following joint research question was formulated: do Dutch-learning infants use language-specific rhythmic patterns for word segmentation or do both Dutch- and Turkish-learning infants use universal cues, such as stress clash and lapse, for word segmentation at 8 months of age? Based on the results from the preference experiments, I predicted that Dutch-learning infants would regard an SW stress pattern as a word-like unit, whereas for the Turkish-learning infants who did not present a language-specific rhythmic preference at 6 months of age, I predicted that they would instead use universal cues such as stress clash (two adjacent stressed syllables) as a discontinuity cue in word segmentation at 8 months of age, although the use of this universal cue is also plausible for Dutch-learning infants at this age.

The results of the predictive segmentation experiment described in Chapter 7 demonstrate that neither Dutch- nor Turkish-learning infants use the WSW stress pattern as a cue to word segmentation, disproving the use of language-specific word stress representations for word segmentation in Dutch-learning infants at this age and in this type of experiment. For Turkish-learning infants a facilitating effect of the WSW stress pattern was neither expected, nor found. Evidence for the use of stress clash as a universal cue to word segmentation was, however, found cross-linguistically as both the Dutch and Turkish language groups segmented the target word when it was embedded in the WWS condition. This finding confirms the use of stress clash as a segmentation strategy across different types of languages at 8 months of age. Moreover, a facilitating effect on word segmentation was also found for the SWW condition in both language groups and it was explained to be caused by an 'edge-effect', also reported on in the literature. This 'edge-effect' is accounted for by the alignment of stressed syllables with phrase boundaries in this particular condition in this word segmentation experiment.

There is also, however, an alternative explanation of these latter results, one invoking a more general salience effect. In both SWW and WWS conditions, the strong syllables were preceded by two weak syllables, also called lapses. This lapse context could have caused the strong syllables to literally 'stand out from the crowd', making it easier to segment the strong target words in these two conditions as compared to in the WSW condition, in which weak and strong syllables alternate rapidly, making the strong target words relatively less salient. The idea of relative salience is also called entropy, or distinctiveness, which was addressed by van Ommen (2016) when examining the 'learnability' of word stress patterns from a distributional perspective by means of computational modeling. She demonstrated that both distributional distinctiveness and the salience of edges improve the likelihood of making correct word segmentations in a corpus in which only phrase boundaries are marked. It thus seems well-motivated to assume infants also use these types of information as a basis in developing a successful word segmentation strategy.

In sum, from the results of the preference experiments, I conclude that a combination of the Iambic-Trochaic Law and the phrasal prominence proposal account for the development of monolingual Dutch- and Turkish-learning infants' rhythmic preferences between 4 and 8 months of age. The segmentation experiment's results reveal that monolingual Dutch- and Turkish-learning infants at 8 months of age do not use language-specific stress cues for word segmentation (yet), but they do use 'edge' stress cues related to perceptual salience and a universal stress clash cue to segmenting target words from an artificial language.

8.1 Suggestions for future research

With regard to future research on the development of language-specific rhythmic preferences, I would suggest studies on older Turkish-learning infants (from 9 months of age) in which they are presented with segmentally varied disyllabic (pseudo-)word stimuli in order to increase the task complexity and prevent early switches to novelty preferences from taking place. The motivation for this suggestion comes from the study on 9-month-old Hebrew-learning infants who demonstrated a native WS preference when presented with segmentally varied low-frequency Hebrew words (Segal & Kishon-Rabin, 2012). It seems plausible to assume that Turkish-learning infants at this age may demonstrate a WS native preference as well, when they are tested in an experiment providing the right conditions for eliciting native, familiarity, preferences.

Concerning the different levels of processing, that is, phrase level versus word level processing, I would also like to suggest that future research should address this distinction more explicitly. This can be done in two ways: either by testing infants at different developmental stages, well before the age at which infants start developing a protolexicon, and well after the age at which infants have developed a substantial lexicon with similar materials (this can be done by employing methods which are suitable for a wider age range and less dependent on cognitive differences between age groups, such as functional Near-Infrared Spectroscopy, fNIRS), or by making the material with which infants are tested at the same age more distinct between the word and the phrase level. The latter option can be realized by comparing a condition in which infants older than 8 months of age are presented with continuous speech without any phrase boundaries to a condition with clear phrase boundaries in a within-subject design. I would predict that younger infants and older infants in the continuous speech condition will demonstrate phrase level rhythmic preferences, while older infants and infants in the segmented condition will demonstrate word level rhythmic preferences. More specifically, I would be very curious to see the results of a rhythmic preference study testing 4-month-old English-learning infants, as I would predict that they show a WS phrase level preference at this age.

Although I believe the word segmentation experiment of the current study has provided interesting results, the complex design of the experiment makes the results difficult to interpret in a conclusive manner. Usually in this type of word segmentation experiment, only two conditions are compared in the test phase, while in the current experiment there were four test conditions. A suggestion for future research would thus be to test the four conditions in a pair-wise between subjects design in which each infant is only presented with two conditions in the test phase, but all possible comparisons between the conditions are still tested in the experiment as a whole. This design would make the outcomes more interpretable. Another suggestion for a future word segmentation experiment is to change the materials by implementing unstressed target words. With this, only the effect of the preceding context is tested, while preventing more global effects such as clash, lapse and aspects of entropy.

In conclusion, and in the 'spirit' of this dissertation, I would like to end with a methodological suggestion for future research. Infant researchers are confronted with the challenge of conducting experiments with children who are, as the Latin origin of the term reveals, 'not speaking' and hence we have to rely on indirect measures of their (linguistic) knowledge, such as the attentional measures used in this study as well as in the many studies mentioned in this dissertation. Not all linguists are necessarily experts on the development of and mechanisms behind infants' attention, which, though understandable is at the same time highly objectionable if we want to be able to interpret and draw reliable conclusions from these measures. The fact that we cannot reliably predict in advance or fully explain afterwards when or why infants demonstrate familiarity or novelty preferences is a thorny issue in the field of infant research in general. Luckily, much work is already being done, for example by Kidd and colleagues (2010; 2012; 2014), aiming to gain more insight into this matter and I highly encourage this type of studies in the field of experimental developmental linguistics for future research.

Bibliography

Abboub, N., Bijeljac-Babic, R., Serres, J. & Nazzi, T. (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology, 132*, 111-120.

Abboub, N., Boll-Avetisyan, N., Bhatara, A., Höhle, B. & Nazzi, T. (2016). An exploration of rhythmic grouping of speech sequences by French- and German-learning infants. *Frontiers in Human Neuroscience, 10*, 292.

Abercrombie, D. (1967). *Elements of General Phonetics*. Aldine Publishing Company.

Adam, G. & Bat-El, O. (2008). Segmental effects on syllable selection: evidence from Hebrew. In A. Gavarró & M. Frietas (Eds.), *Language Acquisition and Development: Proceedings of GALA 2007,* 1-11. Cambridge Scholars Publishing.

Adam, G. & Bat-El, O. (2009). When do universal preferences emerge in language development? The acquisition of Hebrew stress. *Brill's Journal of Afroasiatic Languages and Linguistics*, *1*(1), 255-282.

Adriaans, F. (2006). *PhonotacTools Test version*. Software, Utrecht Institute of Linguistics OTS, Utrecht University.

Adriaans, F. & Kager, R. (2010). Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language, 62*, 311–331.

Allen, G. D. & S. Hawkins (1978). *The Development of Phonological Rhythm.* In Bell & Hooper (Eds.), 173-185.

Allen, G. & Hawkins, S. (1980). Phonological rhythm: Definition and development. *Child Phonology*, *1*, 227-256.

Altvater-Mackensen, N., van der Feest, S. & Fikkert, P. (2014). Asymmetries in early word recognition: the case of stops and fricatives. *Language Learning and Development*, *10*(2), 149-178.

Aslin, R. (2007). What's in a look? *Developmental Science*, *10*(1), 48-53.

Aslin, R., Saffran, J. & Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324.

Baayen, R., Piepenbrock, R. & Gulikers, L. (1995). The CELEX lexical database. University of Pennsylvania, Philadelphia. *Linguistic Data Consortium*.

Bergelson, E. & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T. & Höhle, B. (2013). Native language affects rhythmic grouping of speech. *The Journal of the Acoustical Society of America*, *134*(5), 3828-3843.

Bijeljac-Babic, R., Höhle, B. & Nazzi, T. (2016). Early prosodic acquisition in bilingual infants: The case of the perceptual trochaic bias. *Frontiers in Psychology, 7*, 210.

Bijeljac-Babic, R., Serres, J., Höhle, B. & Nazzi, T. (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLoS ONE, 7(2)*, e30843.

Bion, R., Benavides-Varela, S. & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, *54*(1), 123–140.

Blom, G. (1958). *Statistical Estimates and Transformed Beta-Variables.* John Wiley & Sons.

Boersma, P. & Weenink, D. (2011). *Praat: Doing phonetics by computer.* Software, version 5.2.

Boersma, P. & Weenink, D. (2013). *Praat: Doing phonetics by computer*. Software, version 5.3.

Booij, G. (1995). *The Phonology of Dutch. The Phonology of the World's Languages*. New York.

Bosch, L. & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33-69.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

Cooper, R. & Aslin, R. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584-1595.

Christophe, A., Nespor, M., Teresa Guasti, M. & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science*, *6*(2), 211-220.

Curtin, S., Mintz, T. & Christiansen, M. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233-262.

Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113–121.

Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51–62.

Dember, W. & Earl, R. (1957). Analysis of exploratory, manipulatory, and curiosity behaviors. *Psychological Review*, *64*(2), 91.

Domahs, U., Genc, S., Knaus, J., Wiese, R. & Kabak, B. (2012a). Processing (un-) predictable word stress: ERP evidence from Turkish. *Language and Cognitive Processes, 28*, 335–354.

Dresher, B. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry*, *30*(1), 27-67.

Dresher, B. & Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition*, *34*(2), 137-195.

Dupoux, E., Pallier, C., Sebastián, N. & Mehler, J. (1997). A destressing 'deafness' in French? *Journal of Memory and Language, 36*, 406– 421.

Dupoux, E., Peperkamp, S. & Sebastián-Gallés, N. (2001). A robust method to study stress 'deafness'. *The Journal of the Acoustical Society of America, 110*, 1606–1618.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E. & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

Echols, C. (1987). *A Perceptually-Based Model of Children's First Words.* Doctoral Dissertation, University of Illinois.

Echols, C. (1988). The role of stress, position and intonation in the representation and identification of early words. *Papers and Reports on Child Language Development, 27*, 39-46.

Endress, A. & Hauser, M. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology*, *61*(2), 177-199.

Endress, A. & Mehler, J. (2009b). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language, 60*(3), 351– 367.

Fikkert, P. (1994). *On the Acquisition of Prosodic Structure*. Doctoral Dissertation. Holland Institute of Generative Linguistics, Leiden University.

Frick, J., Colombo, J. & Saxon, T. (1999). Individual and developmental differences in disengagement of fixation in early infancy. *Child Development*, *70*(3), 537-548.

Friederici, A., Friedrich, M. & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*, 1208–1211.

Friederici, A. & Wessels, J. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics*, *54*(3), 287-295.

Gambell, T. & Yang, C. (2006). *Word Segmentation: Quick but not Dirty.* Unpublished manuscript, Yale University.

Giovanelli, G., Sansavini, A. & Farneti, A. (1999). Perception of sound, rhythm and speech from pre-natal to post-natal life. In *Current Issues in Developmental Psychology*, 137–159. Springer.

Gnanadesikan, A. (1995). *Markedness and Faithfulness Constraints in Child Phonology*. Manuscript, University of Massachusetts.

Goedemans, R. & van der Hulst, H. (2009). StressTyp: a database for word accentual patterns in the world's languages. In M. Everaert, S. Musgrave & A. Dimitriadis (Eds.), *The Use of Databases in Cross-Linguistics Studies*, 235–282. Mouton de Gruyter.

Göksel, A. & Kerslake. C. (2005). *Turkish: A Comprehensive Grammar*. Routledge.

Goldstein, H. (2011). *Multilevel Statistical Models.* John Wiley & Sons.

Gout, A. (2001). *Etapes Précoces de l'Acquisition du Lexique.* Unpublished Dissertation, École des Hautes Études en Sciences Sociales.

Goyet, L., Nishibayashi, L. & Nazzi, T. (2013). Early syllabic segmentation of fluent speech by infants acquiring French. *PLoS ONE, 8(11)*, e79646.

Goyet, L., de Schonen, S. & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: An ERP study. *Brain Research, 1332*, 75-89.

Grabe, E. & Low, E. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology,* 7, 515–546.

Grammont, M. (1965). *Traité de Phonétique*. Delagrave.

Güneș, G. (2015). *Deriving Prosodic Structures.* Doctoral Dissertation, University of Groningen.

Gussenhoven, C. (1991). The English Rhythm Rule as an accent deletion rule. *Phonology, 8*, 1–35.

Gussenhoven, C. (2009). Vowel duration, syllable quantity and stress in Dutch. In K. Hanson & S. Inkelas (Eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 181–198. MIT Press.

Halle, M. & Vergnaud, J. (1987). *An Essay on Stress*. MIT Press.

Hammond, M. (1984). *Constraining Metrical Theory: A Modular Theory of Stress and Destressing*. Doctoral Dissertation, University of California at Los Angeles.

Haugen, E. & Joos, M. (1972). Tone and intonation in East Norwegian. In D. Bolinger (Ed.), *Intonation,* 414-436. Penguin.

Hay, J. & Diehl, R. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception & Psychophysics*, *69*(1), 113-122.

Hayes, B. (1980). *A Metrical Theory of Stress Rules*. Doctoral Dissertation, Massachusetts Institute of Technology.

Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry, 15*, 33–74.

Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 429–446.

Hayes, B. (1987). A revised parametric metrical theory. *Proceedings of NELS, 17*(1), 274-189.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press.

Herold, B., Höhle, B., Walch, E., Weber, T. & Obladen, M. (2008). Impaired word stress pattern discrimination in very-low-birthweight infants during the first 6 months of life. *Developmental Medicine & Child Neurology*, *50*(9), 678-683.

Hochberg, J. (1988). First steps in the acquisition of Spanish stress. *Journal of Child Language, 15*, 273–292.

Höhle, B. (2002). *Der Einstieg in die Grammatik: Die Rolle der Phonologie/Syntax-Schnittstelle für Sprachverarbeitung und Spracherwerb.* Unpublished doctoral dissertation, Freie Universität Berlin.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J. & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262-274.

Höhle, B. & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science, 6*, 122–127.

Houston, D., Jusczyk, P., Kuijpers, C., Coolen, R. & Cutler, A. (2000). Cross language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, *7*(3), 504-509.

Hunter, M. & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research 5,* 69-95.

Hyde, B. (2002). A restrictive theory of metrical stress. *Phonology*, *19*(3), 313-359.

Ingram, D. (1974). Phonological rules in young children. *Journal of child language*, *1*(1), 49-64.

Inkelas, S. (1999). Exceptional stress-attracting suffixes in Turkish: representations versus the grammar. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The Prosody-Morphology Interface*, 134–187. Cambridge University Press.

Inkelas, S., Küntay, A., Orgun, C. & Sprouse, R. (2000). Turkish Electronic Living Lexicon (TELL). *Turkic Languages, 4*, 253–275.

Inkelas, S. & Orgun, C. (2003). Turkish stress: A review. *Phonology, 20*, 139–161.

Ipek, C. & Jun, S. (2013). Towards a model of intonational phonology of Turkish: Neutral intonation. In *Proceedings of Meetings on Acoustics, 19*(1), 1-9. Acoustical Society of America.

Ipek, C. & Jun, S. (2014). Distinguishing phrase-final and phrase-medial high tone on finally stressed words in Turkish. In N. Campbell, D. Gibbon & D. Hirst (Eds.), *The Proceedings of the 7th Speech Prosody International Conference*, 393-397.

Johnson, E. & Jusczyk, P. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548-567.

Johnson, E. & Tyler, M. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, *13*(2), 339-345.

Jun, S. & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modeling and technology*, 209–242. Kluwer Academic Publishers.

Jusczyk, P., Cutler, A. & Redanz, N. (1993a). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675–687.

Jusczyk, P., Friederici, A., Wessels, J., Svenkerud, V. & Jusczyk, A. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*(3), 402.

Jusczyk, P., Hohne, E. & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, *61*(8), 1465-1476.

Jusczyk, P., Houston, D. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159–207.

Jusczyk, P. & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics, 23*, 105–109.

Kabak, B. (2011). Turkish vowel harmony. In C. Ewen, E. Hume, M. van Oostendorp & K. Rice (Eds.), *The Blackwell Companion to Phonology*, 2831-2854. Wiley-Blackwell.

Kabak, B. & Revithiadou, A. (2009a). From edgemost to lexical stress: Diachronic paths, typology and representation. *The Linguistic Review, 26*(1), 1-36.

Kabak, B. & Revithiadou, A. (2009b). An interface approach to prosodic word recursion. In J. Grijzenhout & B. Kabak (Eds.), *Phonological domains: Universals and deviations,* 105-133. Mouton de Gruyter.

Kabak, B. & Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology, 18,* 315–360.

Kager, R. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Foris Publications.

Kager, R. (1997). Generalized alignment and morphological parsing. *Rivista di Linguistica, 9*, 245-282.

Kager, R. (2007). Feet and Metrical Stress. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*, 195–227. Cambridge University Press.

Kager, R. & Zonneveld, W. (1986). Schwa, syllables, and extrametricality in Dutch. *The Linguistic Review, 5*, 197–221.

Kamali, B. (2011). *Topics at the PF Interface of Turkish*. Doctoral Dissertation, Harvard University.

van Kampen, A., Parmaksiz, G., van de Vijver, R. & Höhle, B. (2008). Metrical and statistical cues for word segmentation: The use of vowel harmony and word stress as cues to word boundaries by 6- and 9-month-old Turkish learners. In A. Gavarró & M. Frietas (Eds.), *Language Acquisition and Development: Proceedings of GALA 2007*, 313-324. Cambridge Scholars Publishing.

Kaye, J. (1989). *Phonology: A Cognitive View*. Lawrence Erlbaum.

Kidd, C., Piantadosi, S. & Aslin, R. (2010). The Goldilocks Effect: Infants' preference for stimuli that are neither too predictable nor too surprising. In S. Ohlsson (Ed.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 2476-2481. Curran Associates.

Kidd, C., Piantadosi, S. & Aslin, R. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), e36399.

Kidd, C., Piantadosi, S. & Aslin, R. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795-1804.

Konrot, A. (1981). Physical correlates of linguistic stress in Turkish. *University of Essex Language Centre Occasional Papers*, *24*, 26-53.

Kooijman, V., Hagoort, P. & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research, 24(1)*, 109-116.

Kooijman, V., Hagoort, P. & Cutler, A. (2009). Prosodic structure in early word segmentation: ERP evidence from Dutch ten-month-olds. *Infancy, 14(6)*, 591-612.

Kooijman, V., Junge, C., Johnson, E., Hagoort, P. & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Psychology, 4*, 10.3389.

Koopmans-Van Beinum, F. (1980). *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions.* Academische Pers.

Kuijpers, C. T., Coolen, R., Houston, D. & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In: C. Rovee-Collier, L. Lipsitt & H. Hayne (Eds.), *Advances in Infancy Research, 12*, 205-220.

Kurby, C. & Zacks, J. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72-79.

Ladefoged, P. (1975). *A Course in Phonetics*. Harcourt Brace Jovanovich.

Levelt, C. (2012). Perception mirrors production in 14-and 18-month-olds: The case of coda consonants. *Cognition*, *123*(1), 174-179.

Levi, S. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association*, *35*(1), 73-97.

Liberman, M. (1975). *The Intonational System of English*. Doctoral Dissertation, Massachusetts Institute of Technology.

Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249–336.

Mattys, S., Jusczyk, P., Luce, P. & Morgan, J. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38*, 465–494.

McCarthy, J. & Prince, A. (1993). *Generalized Alignment*. Springer.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.

Morgan, J. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language*, *35*(5), 666-688.

Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756-766.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S. & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French:

Emerging evidence for cross-linguistic differences. *Journal of Memory and Language*, *54*(3), 283-299.

Nazzi, T., Jusczyk, P. & Johnson, E. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1-19.

Nazzi, T. & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*(1), 233-243.

Nespor, M., Guasti, M. & Christophe, A. (1996). Selecting word order: the Rhytmic Activation Principle. In U. Kleinhenz (Ed.), *Interfaces in Phonology,* 1 26. Akademie Verlag.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H. & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*, *7*(2), 1-28.

Nespor, M. & Vogel, I. (1986). Prosodic Phonology. *Studies in Generative Grammar, 28*. Foris Publications.

Nespor, M. & Vogel, I. (1989). On clashes and lapses. *Phonology*, *6*(1), 69-116.

Nespor, M., Shukla, M. & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. In C. Ewen, E. Hume, M. van Oostendorp & K. Rice (Eds.), *The Blackwell Companion to Phonology*, 1147–1159. Wiley-Blackwell.

Nishibayashi, L., Goyet, L. & Nazzi, T. (2015). Early speech segmentation in French-learning infants: Monosyllabic words versus embedded syllables. *Language and Speech, 58(3)*, 334-350.

van Ommen, S. (2016*). Listen to the beat: A cross-linguistic perspective on the use of stress in segmentation.* Doctoral Dissertation, Utrecht University. LOT.

van Oostendorp, M. (1995). *Vowel Quality and Phonological Projection.* Doctoral Dissertation, University of Tilburg.

den Os, E. (1988). *Rhythm and Tempo of Dutch and Italian: A Contrastive study*. Dr. Elinkwijk.

Pike, K. (1945). *The Intonation of American English*. University of Michigan Press.

Polka, L. & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. In M. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 1021–1024. Causal Productions.

Polka, L. & Sundara, M. (2012). Word segmentation in monolingual infants acquiring Canadian English and Canadian French: Native language, cross-dialect, and cross-language comparisons. *Infancy*, *17*(2), 198-232.

Polka, L., Sundara, M. & Blue, S. (2002). The role of language experience in word segmentation: A comparison of English, French, and bilingual infants. *The Journal of the Acoustical Society of America*, *111*(5), 2455-2455.

Pons, F. & Bosch, L. (2010). Stress pattern preference in Spanish-learning infants: The role of syllable weight. *Infancy*, *15*(3), 223-245.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry, 14*, 19–100.

Quené, H. & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*, 413–425.

Quené, H. & Koster, M. (1998). Metrical segmentation in Dutch: Vowel quality or stress? *Language and Speech, 41,* 185–201.

Rose, Y. & C. Champdoizeau. (2008). There is no trochaic bias: Acoustic evidence in favor of the Neutral Start Hypothesis. In A. Gavarró & M. Frietas (Eds.), *Language Acquisition and Development: Proceedings of GALA 2007,* 359 – 369. Cambridge Scholars Publishing.

Ramus, F., Dupoux, E. & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In M. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 1021–1024. Causal Productions.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265–292.

Saffran, J., Aslin, R. & Newport, E. (1996). Statistical learning by 8- month-old infants. *Science, 274*, 1926–1928.

Schiering, R. (2007). The phonological basis of linguistic rhythm: Cross-linguistic data and diachronic interpretation. *STUF–Sprachtypologie und Universalienforschung*, *60*(4), 337-359.

Schreuder, M. (2006). *Prosodic processes in language and music*. Doctoral dissertation, Rijksuniversiteit Groningen.

Segal, O. & Kishon-Rabin, L. (2012). Evidence for language-specific influence on the preference of stress patterns in infants learning an iambic language (Hebrew). *Journal of Speech, Language, and Hearing Research, 55*, 1329–1341.

Selkirk, E. (1980). The role of prosodic categories in English word stress. *Linguistic inquiry*, *11*(3), 563-605.

Selkirk, E. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press.

Sezer, E. (1983). On non-final stress in Turkish. *Journal of Turkish Studies, 5*, 61–69.

Shi, R., Cutler, A., Werker, J. & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, *119*(6), EL61-EL67.

Shi, R. & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, *11*(3), 407-413.

Skoruppa, K., Pons, F., Bosch, L., Christophe, A., Cabrol, D. & Peperkamp, S. (2013). The development of word stress processing in French and Spanish infants. *Language Learning and Development*, *9*(1), 88-104.

Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, Alves Limissuri, R. & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, *12*(6), 914-919.

Sluijter, A. & van Heuven, V. (1996a). Acoustic correlates of linguistic stress and accent in Dutch and American English. In H. Bunnell & W. Idsardi (Eds.), *Fourth International Conference on Spoken Language Processing*, 630–633.

Sluijter, A. & van Heuven, V. (1996b). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America, 100*, 2471–2485.

Sluijter, A., van Heuven, V. & Pacilly. (1997). Spectral balance as a cue in the perception of linguistic stress. *The Journal of the Acoustical Society of America, 101*, 503–513.

Smith, N. (1973). *The Acquisition of Phonology*. Cambridge University Press.

Snijders, T. (2011). Multilevel Analysis. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science,* 879-882. Springer.

Swallow, K. & Zacks, J. (2008). Sequences learned without awareness can orient attention during the perception of human activity. *Psychonomic Bulletin & Review*, *15*(1), 116-122.

Swallow, K., Zacks, J. & Abrams, R. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236-257.

Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology, 50*, 86–132.

Thiessen, E., Hill, E. & Saffran, J. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Thiessen, E. & Saffran, J. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Thiessen, E. & Saffran, J. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, *3*(1), 73-100.

Tobii Technology AB. (2006). *Tobii Eye Tracker User Manual*. ClearView analysis software. ©

Topbaş, S. (1997). Phonological acquisition of Turkish children: Implications for phonological disorders. *European Journal of Disorders of Communication*, *32*(4), 377-396.

Toro-Soto, J., Rodriguez-Fornells, A. & Sebastián-Gallés, N. (2007). Stress placement and word segmentation by Spanish speakers. *Psicologica, 28*, 167–176.

Trommelen, M. & Zonneveld, W. (1989). *Klemtoon en Metrische Fonologie*. Dick Coutinho.

Trubetzkoy, N. (1939/1969). *Principles of Phonology*. University of California Press. (Translated by C. Baltaxe.)

Tyler, M. & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America, 126*, 367–376.

Underhill, R. (1976). *Turkish Grammar*. MIT Press.

Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In J. Sundberg, L. Nord & R. Carlson (Eds.) *Music, Language, Speech and Brain*, 108-120. Macmillan Education.

Veenker, T. (2014). *The Zep Experimental Control Application*. Software, version 1.6. Utrecht Institute of Linguistics OTS, Utrecht University.

Vroomen, J. & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 98–108.

Wass, S., Smith, T. & Johnson, M. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229-250.

White, L., Mattys, S. & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language, 66*, 665–679.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology, 14*, 1-66.

Woodrow, H. (1911). The role of pitch in rhythm. *Psychological Review, 18*, 54-77.

Woodrow, H. (1951). Time perception. In S. Stevens (Ed.), *Handbook of Experimental Psychology*, 1224-1236. John Wiley & Sons.

Yoshida, K., Iversen, J., Patel, A., Mazuka, R., Nito, H., Gervain, J. & Werker, J. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition*, *115*(2), 356-361.
van Zonneveld, R. (1985). Word rhythm and the Janus syllable. In H. van der Hulst & N. Smith (Eds.), *Advances in Nonlinear Phonology*, 133–142. Foris Publications.

Samenvatting in het Nederlands

Hoofdstuk 1
Algemene inleiding

Volwassen sprekers hebben onbewuste kennis van de klankregels van hun moedertaal of talen. Deze klankregels zijn onderdeel van de fonologie van een taal en omvatten ook kennis van de ritmische structuur van een taal. Deze ritmische structuur, zoals beschreven in de metrische fonologie, bepaalt onder andere welke lettergrepen in een woord klemtoon krijgen. In veel talen krijgen de lettergrepen aan de 'woordranden' klemtoon, zoals de eerste of (voor)laatste lettergreep. De verwachting is dat het kenmerk van woordklemtoon om aan de woordranden te vallen een fonologische eigenschap is die luisteraars helpt bij het opdelen van lopende spraak in woorden. Hoewel zowel begin- als eindklemtoon voorkomt in de talen van de wereld, werd er tot voor kort bijna uitsluitend bewijs voor deze 'metrische segmentatiehypothese' geleverd door talen met beginklemtoon. Nu we weten dat volwassen sprekers van talen met eindklemtoon ook gebruik kunnen maken van deze segmentatiestrategie rees de vraag hoe deze strategie verworven wordt door baby's. Het onderzoek beschreven in deze dissertatie neemt daarom een taal-overschrijdend perspectief in om te onderzoeken wanneer leerders van een taal met eindklemtoon dit patroon gaan gebruiken om woorden te ontdekken in lopende spraak.

Om deze kwestie te onderzoeken is een serie experimenten uitgevoerd met als doel de volgende vragen te beantwoorden: (i) wanneer leren baby's het ritmische patroon van hun moedertaal, waarbij de onderzochte talen verschillen in klemtoonpatroon, namelijk begin- versus eindklemtoon, en (ii) gebruiken baby's het klemtoonpatroon van hun moedertaal als hulp bij het segmenteren van woorden of gebruiken ze hierbij geen taal-specifieke, maar juist 'universele' aanwijzingen? Voordat deze vragen in detail onderzocht worden, wordt eerst in hoofdstuk 2 een overzicht gegeven van de bestaande literatuur over ritmische structuur, ritmische discriminatie, ritmische voorkeuren en spraaksegmentatie door baby's. Hierna worden de fonologische eigenschappen van de onderzochte talen, Nederlands en Turks, uiteengezet in hoofdstuk 3. Om onderzoeksvraag (i) te beantwoorden werden de twee experimenten beschreven in hoofdstuk 4 en 5 uitgevoerd en de resultaten van deze experimenten worden gezamenlijk besproken in hoofdstuk 6.

De tweede onderzoeksvraag (ii) wordt beantwoord door middel van het segmentatie-experiment beschreven in hoofdstuk 7. Tot slot worden de resultaten van alle experimenten afgezet tegen de hypotheses die beschreven worden in het literatuuroverzicht en de hypotheses die worden voorgesteld naar aanleiding van de resultaten. Het laatste hoofdstuk, hoofdstuk 8, geeft daarnaast de algemene conclusies van het huidige onderzoek en doet suggesties voor toekomstig onderzoek.

Hoofdstuk 2
Literatuuroverzicht

Ritmische structuur

In de 'metrische fonologie' wordt de ritmische structuur van talen op woordniveau beschreven in termen van 'metrische voeten': combinaties van twee lettergrepen die ofwel beginklemtoon (een trocheïsch patroon), ofwel eindklemtoon (een jambisch patroon) hebben. Volgens de 'jambisch-trocheïsche wet' worden deze patronen ook verschillend uitgesproken, waarbij de jambe vooral verschilt in de duur van de lettergrepen (kort-lang), terwijl de trochee vooral verschilt in toonhoogte tussen de lettergrepen (hoog-laag). Uit meerdere onderzoeken is gebleken dat mensen op basis van duur- of toonhoogteverschillen inderdaad, respectievelijk, jamben of trocheeën waarnemen. We weten ook dat baby's vroeg in hun ontwikkeling een grotere gevoeligheid hebben voor toonhoogteverschillen dan voor duurverschillen en dat baby's daarom wellicht een voorkeur zouden hebben voor trocheeën, onafhankelijk van de taal die ze leren. Deze verwachting wordt beschreven in de 'universele trocheïsche voorkeurshypothese'. Daarnaast weten we dat talen verschillen in hun 'timing': in sommige talen, zoals Romaanse talen, zijn alle lettergrepen van gelijke duur, terwijl in andere talen, zoals Germaanse talen, alle metrische voeten gelijke duur hebben, maar lettergrepen juist verschillen in duur afhankelijk van of ze beklemtoond zijn of niet. De 'ritmische klassenhypothese' voorspelt dat alleen voor talen in de laatste klasse klemtoon belangrijk is voor spraaksegmentatie.

Ritmische discriminatie en voorkeuren

Ritmische discriminatie is het horen van verschillen tussen talen uit de hierboven beschreven ritmische klassen en het horen van het verschil tussen jamben en trocheeën. Uit onderzoek blijkt dat pasgeboren baby's al in staat zijn het verschil te horen tussen talen uit verschillende ritmische klassen en dat ze rond vier maanden ook het verschil kunnen horen tussen jamben en trocheeën, onafhankelijk van de taal die de baby's leren. Het ontwikkelen van ritmische voorkeuren, het laten zien van een voorkeur voor het klemtoonpatroon van de moedertaal, lijkt echter wel taal-afhankelijk te zijn. Baby's die Germaanse talen leren, zoals Engels en Duits, laten met zes maanden een voorkeur zien voor het klemtoonpatroon van hun moedertaal, terwijl baby's die Frans leren geen voorkeur laten zien. De reden voor dit verschil wordt in deze dissertatie eerst gezocht in methodologisch verschillen tussen de beschreven onderzoeken.

Methodologische kwesties

Sommige experimenten uit de literatuur gebruiken namelijk één woord waarbinnen alleen de plek van de klemtoon verschilt als stimuli (het spraakmateriaal waarnaar baby's luisteren tijdens een experiment), terwijl andere experimenten verschillende woorden gebruiken. Het blijkt voor baby's gemakkelijker om klemtoonpatronen te discrimineren of een voorkeur te laten zien wanneer slechts één woord gebruikt wordt. Hiermee wordt daarom rekening gehouden in het huidige onderzoek. Daarnaast verschilde het spraakmateriaal ook in het feit of de baby's luisterden naar een spreker van hun moedertaal of niet. Het kan dus zijn dat de spreker van de stimuli beïnvloedt of baby's een voorkeur laten zien voor het klemtoonpatroon van hun moedertaal. Dit idee wordt later in de dissertatie ook beschreven in de 'moedertaal-afhankelijkheidshypothese'. Verder verschilden de experimenten ook in de manier waarop klemtoon was aangebracht in de stimuli. Soms werd dit gedaan door middel van natuurlijke spraak en soms middels akoestische manipulatie. Er wordt benadrukt dat de specifieke akoestische kenmerken die klemtoon vormen, zoals duur, toonhoogte en luidheid, kunnen beïnvloeden of baby's een voorkeur laten zien voor hun moedertaalklemtoonpatroon of niet.

Spraaksegmentatie

Woordsegmentatie-experimenten met baby's hebben laten zien dat ze verschillende aanwijzingen gebruiken voor het ontdekken van woorden in lopende spraak, waarvan klemtoon er één is. Baby's die Engels leren gebruiken al met 7,5 maand het klemtoonpatroon van hun moedertaal om woorden te segmenteren. Op basis van deze bevinding werd de 'metrische segmentatiehypothese' opgesteld, maar het was onbekend hoe universeel of taal-specifiek deze hypothese moet worden geïnterpreteerd. In het Frans wordt namelijk geen bewijs gevonden voor het gebruik van klemtoon in woordsegmentatie.

Hoofdstuk 3
Talen: Nederlands en Turks

Als reden voor het feit dat in het Frans noch een ritmische voorkeur noch het gebruik van klemtoon in segmentatie wordt gevonden, wordt aangedragen dat het Frans eigenlijk geen klemtoon zou hebben op woordniveau, maar alleen op frase- of zinsniveau. Er wordt daarom gesteld dat het noodzakelijk is een taal te onderzoeken die net als het Frans ook eindklemtoon heeft, maar dan daadwerkelijk op woordniveau. Daarom worden als talen voor het huidige onderzoek Nederlands, qua klemtoon vergelijkbaar met Engels en Duits, en Turks, dat eindklemtoon op woordniveau heeft, gekozen.

Nederlands

Nederlands en Turks vertonen geen honderd procent overlap in de fonemen (klanken) die de talen gebruiken. Het is daarom van belang in het huidige onderzoek alleen te werken met fonemen die wel in beide talen voorkomen. Daarnaast is er in het Nederlands invloed van de lettergreepstructuur op de plek van de klemtoon in woorden: lettergrepen die uit meer fonemen bestaan trekken namelijk klemtoon aan. Het is in het huidige onderzoek dus belangrijk lettergrepen te kiezen die uit evenveel fonemen bestaan, zodat beide lettergrepen klemtoon kunnen krijgen en zo kan worden afgewisseld tussen begin- en eindklemtoon. Verder behoort het Nederlands tot de Germaanse talen en daarmee ook tot de ritmische klasse waarin metrische voeten voorkomen. De meest voorkomende voet in het Nederlands is de trochee, ofwel beginklemtoon op woordniveau. Op frase- of zinsniveau kent het Nederlands echter zowel begin- als eindklemtoon door het gebruik van twee verschillende woordvolgorden.

Turks

In het Turks is iets bijzonders aan de hand met klinkers, ofwel vocalen. Binnen een woord mogen de klinkers namelijk niet wisselen van soort (in het Turks is dat vooral het geval tussen voor- en achterklinkers, maar ook tussen ronde en niet-geronde klinkers). Dit fenomeen heet vocaalharmonie. Wanneer er wel gewisseld wordt tussen klinkers is dit dus een heel duidelijke aanwijzing voor een woordgrens en daarom worden in het huidige onderzoek alleen ‘harmonische’ stimuli gebruikt. Verder kent het Turks een aantal lettergrepen met een speciaal kenmerk, ook wel morfemen genoemd. Deze morfemen zorgen ervoor dat de klemtoon niet op de laatste lettergreep ligt, maar op de voorlaatste lettergreep. Deze lettergrepen worden dus vermeden in het huidige onderzoek. Het Turks behoort tot de ritmische klasse van talen waarin geen metrische voeten voorkomen en lijkt hiermee dus op het Frans. Het Turks heeft in tegenstelling tot het Frans wel eindklemtoon op woordniveau. Op frase- of zinsniveau kent het Turks echter een ander patroon dat vanwege de woordvolgorde overwegend beginklemtoon betreft.

Hoofdstuk 4
Ritmische voorkeuren

Om de eerste onderzoeksvraag te beantwoorden werd een voorkeursexperiment ontwikkeld waarmee zowel Nederlands- als Turks-lerende baby's getest konden worden. Eerst werden de Nederlands-lerende baby's getest en daarna de Turks-lerende baby's.

Experiment 1: methode en resultaten
Aan het eerste experiment namen 102 Nederlands-lerende baby's van tussen de vier en acht maanden oud deel. Ze luisterden tijdens het experiment naar een verzonnen woord (een woord dat in zowel het Nederlands als het Turks geen betekenis heeft) dat afwisselend of begin- of eindklemtoon had. Het woord werd uitgesproken door een moedertaalspreker van het Spaans. Door middel van oogbewegingsapparatuur kon de aandacht van de kinderen tijdens het luisteren naar de verschillende klemtoonpatronen worden gemeten. Uit de resultaten van het experiment blijkt dat Nederlands-lerende baby's een voorkeur laten zien voor het woord met beginklemtoon, wat overeenkomt met het klemtoonpatroon van hun moedertaal. Deze voorkeur is het sterkst wanneer de baby's zes maanden oud zijn.

Experiment 2: methode en resultaten
Aan het tweede experiment namen 92 Turks-lerende baby's van tussen de vier en acht maanden oud deel. Ze luisterden naar dezelfde stimuli als de Nederlandse-lerende baby's en hun aandacht werd op dezelfde manier gemeten. De resultaten laten verrassend genoeg zien dat ook Turks-lerende baby's een voorkeur laten zien voor beginklemtoon, hoewel dit niet het woordklemtoonpatroon van hun moedertaal is. Deze voorkeur is het sterkst met vier maanden oud.

Discussie en conclusie
Er wordt beargumenteerd dat het feit dat beide taalgroepen een voorkeur voor het trocheïsche (beginklemtoon) patroon laten zien de 'universele trocheïsche voorkeur', zoals besproken in de inleiding, reflecteert. Er wordt echter een

aanvullende hypothese voorgesteld waarin wordt gesteld dat geen van beide taalgroepen een taal-specifieke voorkeur laten zien, omdat alle baby's luisteren naar een niet-moedertaalspreker. Deze 'moedertaal-afhankelijkheidshypothese' voorspelt dat taal-specifieke voorkeuren alleen kunnen worden uitgelokt wanneer baby's luisteren naar een spreker van hun moedertaal. Daarom wordt in een vervolgexperiment gewerkt met een tweetalige spreker van zowel het Nederlands als het Turks en wordt de aard van de gevonden 'universele trocheïsche voorkeur' nader onderzocht.

Hoofdstuk 5
De universele trocheïsche voorkeur

In dit vervolgexperiment luisterden 48 Nederlands-lerende baby's en 48 Turks-lerende baby's met een leeftijd tussen de zes en acht maanden opnieuw naar een verzonnen woord dat alleen verschilde in de plek van de klemtoon. Dit keer was het woord voor de Nederlands-lerende baby's uitgesproken in Nederlandse modus en voor de Turks-lerende baby's in Turkse modus. De aandacht van de baby's werd weer gemeten met behulp van oogbewegingsapparatuur. Uit de resultaten blijkt opnieuw dat er geen verschil in voorkeur is tussen de taalgroepen, waarmee de uitkomsten van het eerste experiment gerepliceerd zijn, maar waarmee geen bewijs is gevonden voor de 'moedertaal-afhankelijkheidshypothese'. Om de uitkomsten van beide experimenten zorgvuldig te interpreteren worden de resultaten in een interim-discussie in hoofdstuk 6 afgezet tegen de verschillende hypotheses die tot nu toe de revue zijn gepasseerd.

Hoofdstuk 6
Interim-discussie

De eerste hypothese die wordt besproken is de 'ritmische klassenhypothese'. In combinatie met het 'ritmische activatievoorstel' voorspelt deze hypothese dat alleen baby's die een taal leren met klemtoon-'timing', zoals het Nederlands, ritmische voorkeuren laten zien. Deze hypothese kan echter niet verklaren waarom er ook in de Turks-lerende groep een ritmische voorkeur wordt gevonden, terwijl het Turks lettergreep-'timing' heeft. Als alternatief wordt vervolgens de 'woordniveau klemtoonhypothese' besproken die stelt dat onafhankelijk van ritmische klasse alle baby's die een taal leren met klemtoon op woordniveau een voorkeur zouden moeten laten zien voor het klemtoonpatroon van hun moedertaal. Deze hypothese is echter ook niet toereikend, omdat de Turks-lerende baby's geen voorkeur laten zien voor het eindklemtoonpatroon van hun moedertaal, maar juist voor het tegenovergestelde beginklemtoonpatroon. Daarop wordt er teruggekeerd naar de 'universele trocheïsche voorkeurshypothese', maar wordt uiteindelijk geconcludeerd dat ook deze hypothese niet alle resultaten juist voorspelt: in de Nederlands-lerende groep wordt deze voorkeur namelijk niet al vanaf de jongste leeftijd van vier maanden gevonden, wat de hypothese wel zou voorspellen. Zoals reeds besproken is er ook geen bewijs gevonden voor de voorgestelde 'moedertaal-afhankelijkheidshypothese'.

Het enige voorstel dat de gevonden resultaten zou kunnen verklaren is een combinatie van de 'jambisch-trocheïsche wet' en het 'frase-prominentievoorstel' waarin de akoestische eigenschappen van uitingen op frase- of zinsniveau worden gerelateerd aan de woordvolgorde binnen uitingen. In het Nederlands komen namelijk twee verschillende woordvolgorden voor die elk hun eigen klemtoonpatroon op frase- of zinsniveau kennen: zowel begin als eindklemtoon komen op zinsniveau dus voor in het Nederlands. In het Turks komt echter de woordvolgorde die gemarkeerd wordt door beginklemtoon op frase- of zinsniveau overwegend voor. Gegeven het feit dat baby's voor de leeftijd van zes maanden nog geen woordenschat hebben opgebouwd, lijkt het plausibel dat ze voor die leeftijd vooral aandacht hebben voor de ritmische structuur van frases of zinnen en pas later overgaan op het leren van de ritmische structuur van woorden. Deze conclusie heeft gevolgen voor de verwachte resultaten van het laatste experiment waarin het gebruik van klemtoon voor woordsegmentatie wordt onderzocht.

Hoofdstuk 7
Het gebruik van klemtoon voor woordsegmentatie

Methode

Met dit experiment wordt onderzocht of en hoe Nederlands-lerende en Turks-lerende baby's klemtoon gebruiken voor het segmenteren van woorden uit lopende spraak. Er namen 24 Nederlands-lerende en 24 Turks-lerende baby's van 8 maanden oud deel aan het experiment. Tijdens het experiment konden de baby's 'woorden' ontdekken in een kunstmatige taal. Het experiment bestond daarom uit twee fasen: een leerfase en een testfase. In de leerfase luisterden de baby's naar drie passages van zes 'zinnen' waarin per passage in elke zin een uniek eenlettergrepig 'woord' verstopt zat. Elk passage kende zijn eigen klemtoonpatroon dat altijd vooraf ging aan hetzelfde, ook beklemtoonde, doelwoord. In totaal kon elke baby dus drie woorden leren die elk gekoppeld waren aan een specifiek klemtoonpatroon. Er waren drie verschillende klemtoonpatronen in de leerfase: sterk-zwak-zwak, zwak-sterk-zwak en zwak-zwak-sterk. In de testfase kregen de baby's de drie 'woorden' uit de leerfase te horen, plus een nieuw 'woord' dat niet in de leerfase voorkwam. In de testfase werd opnieuw de aandacht van de baby's gemeten met oogbewegingsapparatuur.

Resultaten

De resultaten laten wederom geen verschil tussen de taalgroepen zien: de Nederlands-lerende en de Turks-lerende baby's gebruiken dus dezelfde klemtoonpatronen als hulp voor het segmenteren van woorden. De woorden uit de passage waarin de 'zin' soms begon met een beklemtoonde lettergreep (het sterk-zwak-zwak patroon) worden als bekend beschouwd in de testfase. Dit blijkt uit een gebrek aan interesse voor deze woorden in de testfase en betekent dat deze woorden gesegmenteerd zijn in de leerfase. Daarnaast worden de woorden uit de passage waarin een beklemtoonde lettergreep direct vooraf ging aan het eveneens beklemtoonde doelwoord (het zwak-zwak-sterk patroon) ook als bekend beschouwd in de testfase. Dit blijkt wederom uit de weinige aandacht voor deze woorden in de testfase en betekent dus dat ook deze woorden gesegmenteerd zijn in de leerfase. De woorden die gekoppeld waren aan het zwak-sterk-zwak patroon kregen in de testfase evenveel aandacht als de nieuwe woorden die helemaal niet voorkwamen in de leerfase. Hieruit blijkt dat de woorden uit deze passage even onbekend waren voor de baby's als de nieuwe woorden, wat betekent dat ze niet gesegmenteerd zijn in de leerfase.

Discussie en conclusie

Aangezien er geen verschil wordt gevonden in hoe de taalgroepen klemtoon gebruikten als hulp voor woordsegmentatie, worden de gevonden segmentatiestrategieën niet als taal-specifieke, maar juist als 'universele' strategieën geïnterpreteerd. In het geval van het sterk-zwak-zwak-patroon gebruiken de baby's een opvallend begin van de zin om het eveneens opvallende doelwoord te segmenteren als het begin van iets nieuws. In het geval van het zwak-zwak-sterk patroon gebruiken ze het feit dat een 'klemtoonbotsing' (twee aangrenzende klemtonen) bijna nooit binnen één woord voorkomt om het doelwoord te segmenteren.

Hoofdstuk 8
Algemene discussie en conclusie

Concluderend kan gesteld worden dat de resultaten van de voorkeursexperimenten met Nederlands-lerende en Turks-lerende baby's tussen 4 en 8 maanden het best verklaard kunnen worden door een combinatie van de 'jambisch-trocheïsche wet' en het 'frase-prominentie voorstel'. Deze verklaring benadrukt dat baby's eerst op frase- of zinsniveau de ritmische structuur van hun moedertaal ontdekken, voordat ze beginnen met het leren van het klemtoonpatroon op woordniveau. De resultaten van het segmentatie-experiment met Nederlands-lerende en Turks-lerende baby's van 8 maanden oud laten zien dat ze (nog) geen taal-specifieke klemtoonpatronen gebruiken voor woordsegmentatie. Ze gebruiken wel de opvallendheid van klemtoon aan het begin van een zin en het feit dat twee aangrenzende klemtonen niet binnen één woord kunnen voorkomen om doelwoorden te segmenteren uit een kunstmatige taal.

Suggesties voor toekomstig onderzoek

Wat betreft toekomstig onderzoek wordt voorgesteld om dit expliciet te richten op het onderscheid tussen het verwerken van spraak op woord- en zinsniveau. Deze overgang lijkt gerelateerd te zijn aan hoever baby's zijn in hun woordenschatontwikkeling, maar voor deze aanname moet het echte bewijs nog worden geleverd in de toekomst.

Curriculum Vitae

Brigitta Keij was born on June 30, 1986 in Voorburg, the Netherlands. Her pre-university education focused on modern foreign languages, but also included the study of biology. After obtaining her VWO-diploma in 2004, she studied at Utrecht University in the Netherlands from 2004 until 2008. In 2008, she graduated with a BA in Language and Culture Studies, which included a pre-master programme in Language and Speech Pathology at the Radboud University Nijmegen, the Netherlands, and subsequently enrolled in the master programme at the same university. In 2009, she obtained her MA degree in Language and Speech Pathology (bene meritum). Brigitta wrote her master thesis on the acquisition of grammatical gender and the definite determiner in Dutch by monolingual and bilingual children with and without a developmental language disorder. The results of this thesis were published in the international journal *Linguistic Approaches to Bilingualism.* During her second master programme in Clinical Linguistics from 2009 until 2011, she spent a trimester at the University of Milano-Bicocca in Italy, a trimester at the University of Groningen, the Netherlands, and a trimester at the University of Potsdam in Germany. In 2011, she concluded her research master and obtained her MSc degree in Clinical Linguistics ('good').

This dissertation is the result of the work Brigitta carried out as a PhD candidate between 2011 and 2014 at the Utrecht Institute of Linguistics OTS of Utrecht University. During that period, she collaborated with Koç University in Istanbul, Turkey, and presented her work at many international conferences and workshops as well as at lab meetings of different universities. Part of the results have been published in the volume *Dimensions of Phonological Stress* by Cambridge University Press. She also worked at the University of Barcelona in Catalonia, Spain, during two short visits in the summers of 2015 and 2016. Since 2015, Brigitta works as a lecturer and junior assistant professor in Linguistics at Utrecht University, continuing her research on language acquisition.